FORM PTO-1390 (Modified) (REV 11-2000) U.S DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: 000487.00012

U S APPLICATION NO (IF KNOWN, SEE 37 CFR): 10/088417

INTERNATIONAL APPLICATION NO	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/GB00/03576	18 September 2000	17 September 1999

TITLE OF INVENTION
PROTEIN STRUCTURES AND PROTEIN FIBRES

APPLICANT(S) FOR DO/EO/US
WOOLFSON, Derek N.; WALSHAW, John; PANDYA, Maya J.; and COLYER, John

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (24) indicated below.
4. ☐ The US has been elected by the expiration of 19 months from the priority date (Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371 (c) (2))
   a. ☒ is attached hereto (required only if not communicated by the International Bureau).
   b. ☐ has been communicated by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
   a. ☐ is attached hereto.
   b. ☐ has been previously submitted under 35 U.S.C. 154(d)(4).
7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☐ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).
11. ☒ A copy of the International Preliminary Examination Report (PCT/IPEA/409).
12. ☒ A copy of the International Search Report (PCT/ISA/210).

Items 13 to 20 below concern document(s) or information included:

13. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
14. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included
15. ☐ A **FIRST** preliminary amendment.
16. ☐ A **SECOND** or **SUBSEQUENT** preliminary amendment.
17. ☐ A substitute specification
18. ☐ A change of power of attorney and/or address letter
19. ☐ A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U S.C. 1.821 - 1 825
20. ☐ A second copy of the published international application under 35 U S.C 154(d)(4).
21. ☐ A second copy of the English language translation of the international application under 35 U S C 154(d)(4)
22. ☐ Certificate of Mailing by Express Mail
23. ☒ Other items or information:

PTO-1449 w/5 references

U.S. APPLICATION NO (IF KNOWN, SEE 37 CFR)	INTERNATIONAL APPLICATION NO	ATTORNEY'S DOCKET NUMBER
10/088417	PCT/GB00/03576	000487.00012

24. The following fees are submitted..		CALCULATIONS PTO USE ONLY
BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5))		
☐ Neither international preliminary examination fee (37 CFR 1 482) nor international search fee (37 CFR 1 445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO	$1040.00	
☒ International preliminary examination fee (37 CFR 1 482) not paid to USPTO but International Search Report prepared by the EPO or JPO . .	$890.00	
☐ International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1 445(a)(2)) paid to USPTO	$740.00	
☐ International preliminary examination fee (37 CFR 1 482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4)	$710.00	
☐ International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4)	$100.00	
ENTER APPROPRIATE BASIC FEE AMOUNT =		$890.00
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☒ 30 months from the earliest claimed priority date (37 CFR 1 492 (c)).		$130.00

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	
Total claims	35 - 20 =	15	x $18.00	$270.00
Independent claims	2 - 3 =	0	x $84.00	$0.00
Multiple Dependent Claims **(check if applicable)**			☒	$280.00

TOTAL OF ABOVE CALCULATIONS =	$1,570.00
☐ Applicant claims small entity status. See 37 CFR 1 27). The fees indicated above are reduced by 1/2	$0.00
SUBTOTAL =	$1,570.00
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (f)). +	$0.00
TOTAL NATIONAL FEE =	$1,570.00
Fee for recording the enclosed assignment (37 CFR 1 21(h)) The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3 31) **(check if applicable).** ☐	$0.00
TOTAL FEES ENCLOSED =	$1,570.00
Amount to be: refunded	$
charged	$

a. ☐ A check in the amount of _____ to cover the above fees is enclosed.

b. ☒ Please charge my Deposit Account No **19-0733** in the amount of **$1,570.00** to cover the above fees. A duplicate copy of this sheet is enclosed.

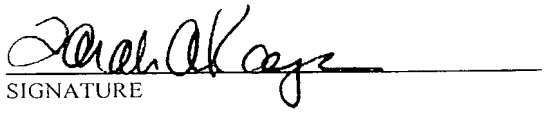
c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. **19-0733** A duplicate copy of this sheet is enclosed

d. ☐ Fees are to be charged to a credit card **WARNING:** Information on this form may become public. **Credit card information should not be included on this form.** Provide credit card information and authorization on PTO-2038

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO

KAGAN, Sarah A.
BANNER & WITCOFF, LTD.
1001 G Street, N.W.
Eleventh Floor
Washington, D.C. 20001-4597



SIGNATURE

Sarah A. Kagan
NAME

32,141
REGISTRATION NUMBER

18 March 2002
DATE

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of	)	Group Art Unit: TBA
Derek WOOLFSON et al.	)	Examiner: TBA
Serial No. 10/088,417	)	
Filed: March 18, 2002	)	Atty. Docket No. 000487.00012

For: PROTEIN STRUCTURES AND PROTEIN FIBRES

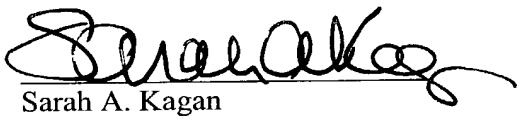
SUBMISSION OF SEQUENCE LISTING

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In response to the Notice to File Missing Requirements mailed May 21, 2002, in the above-identified application, Applicants submit a computer readable form (CRF). The content of the CRF and the paper copy submitted herewith are believed to be the same and to add no new matter. It is believed that no fee is due. However, if such a fee is deemed necessary, please charge our Deposit Account No. 19-0733.

Respectfully submitted,

By: 
Sarah A. Kagan
Registration No. 32,141

Dated: July 22, 2002

Eleventh Floor
1001 G Street, N.W.
Washington, D.C. 20001-4597
(202) 508-9100

#5

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of	)	Group Art Unit: TBA
	)	
Derek WOOLFSON et al.	)	Examiner: TBA
	)	
Serial No. 10/088,417	)	
	)	
Filed: March 18, 2002	)	Atty. Docket No. 000487.00012

For: PROTEIN STRUCTURES AND PROTEIN FIBRES

PRELIMINARY AMENDMENT

Assistant Commissioner of Patents
Washington, D. C. 20231

Sir:

Preliminarily to the examination of the above-identified application, kindly amend the application as follows:

IN THE SPECIFICATION:

Page 1, after the title, insert the following paragraph:

This national phase Application of PCT/GB00/03576 filed September 18, 2000 was published under PCT Article 21(12) in English and claims the priority of GB 9922013.9.

REMARKS

The amendment to the specification is made in accordance with 35 U.S.C. 119, 37 C.F.R. 1.78 and 37 C.F.R. 1.55. No new matter has been entered. Entry is requested.

Respectfully submitted,

Dated: July 18, 2002

Sarah A. Kagan
Registration No. 32,141

BANNER & WITCOFF, LTD.
1001 G Street, N.W.
Eleventh Floor
Washington, D.C. 20001
TEL: (202) 508-9100
FAX: (202) 508-9299

478651-1

#6

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of	)	Group Art Unit: TBA
	)	
Derek WOOLFSON et al.	)	Examiner: TBA
	)	
Serial No. 10/088,417	)	
	)	
Filed: March 18, 2002	)	Atty. Docket No. 000487.00012

For: PROTEIN STRUCTURES AND PROTEIN FIBRES

SECOND PRELIMINARY AMENDMENT

Assistant Commissioner of Patents
Washington, D. C. 20231

Sir:

Preliminarily to the examination of the above-identified application, kindly amend the application as follows:

IN THE CLAIMS:

16. (Amended) A protein structure according to any preceding claim in which the first and second peptide monomer units have the sequence:

a) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-pl; SEQ ID NO: 1) and

b) KIRALKAKNAHLKQEIAALEQEIAALEQ (SAF-p2; SEQ ID NO: 2) respectively; or

c) KIAALKQKIAALKQEIDALEYENDALEQ (SAF-plA; SEQ ID NO: 3) and

d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 4) respectively; or

e) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-pIC; SEQ ID NO: 1) and

f) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 4) respectively.

17. (Amended) A peptide monomer unit for use in preparing a protein structure the peptide monomer unit having an amino acid sequence selected from:

a) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-pl; SEQ ID NO: 1);

b) KIRALKAKNAHLKQEIAALEQEIAALEQ (SAF-p2; SEQ ID NO: 2);

c) KIAALKQKIAALKQEIDALEYENDALEQ (SAF-p1A; SEQ ID NO: 3);

d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 5);

e) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C; SEQ ID NO: 1) ; and

d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 5).

IN THE SPECIFICATION

At page 4, paragraph 3, substitute the following paragraph:

In a preferred protein structure, the first and second peptide monomer units have the following sequences:

a) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C; SEQ ID NO: 1) and
b) KIRALKAKNAHLKQEIAALEQEIAALEQ (SAF-p2D; SEQ ID NO: 2) respectively; or
c) KIAALKQKIAALKQEIDALEYENDALEQ (SAF-p1A; SEQ ID NO: 3) and
d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 4) respectively; or
e) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C; SEQ ID NO: 1) and
f) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 4) respectively.

At page 9, table legend, substitute the following legend:

*=Chemical capping = CH_3CO at the N terminus and NH_2 at the C terminus. The sequences in the table are SEQ ID NO: 5, 6, 7, 8, 9, 10, 10, 11, and 12, respectively.

At page 7, paragraph 5, substitute the following paragraph:

Fig. 1A and Fig. 1B illustrate the design (Fig. 1A) and the sequences (Fig. 1B; SEQ ID NOs: 15 and 16) of self-assembling fibre (SAF) peptide monomers of the invention.

At page 8, paragraph 2, substitute the following paragraph:

Fig. 8 shows amino acid sequences (SEQ ID NOs: 17 and 18) designed to form blunt-ended heterodimers.

At page 12, paragraph 1, substitute the following paragraph:

In addition and as a control, the SAF-p I c sequence was permuted (N- and C-terminal halves were swapped) to produce peptide SAF-p3:

E IDALEYE NDALEQK IAALKQK IASLKQ (SEQ ID NO: 13.)

This design should combine with SAF-p2D to form a blunt-ended structure, which should not form fibres.

At page 12, paragraph 2, substitute the following paragraph:

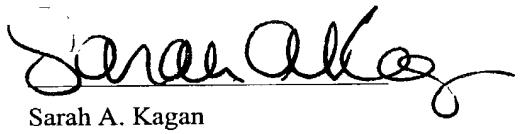
A model of the three-dimensional structure of the designed protein fibre resulting from the assembly of SAF-pl and SAF-p2 was made from the minimised structure of a model coiled-coil 35-mer, (LAALAAA)s (SEQ ID NO: 14), which was generated using Crick's Equation and had an ideally packed interface (G. Offer and R. Sessions, J. Mol. Biol. 249, 967 (1995)). Copies of the 35-mer were superimposed with an overlap of one heptad repeat to extend the structural template, and the backbone was rejoined after removal of overlapping°segments. Residues in the two-stranded template were replaced with the sequences of the SAF peptides, staggered relative to each other by two heptad repeats according to the alignment in. Fig. 1B. The structure was soaked in a 5 A layer of water and energy minimised until the average absolute derivative of coordinates with respect to energy fell below 0.01 kcal A-'. The structure was built and visualized using insight II 97.0 (Molecular Simulations Inc.), and was energy-minimized using Discover 2.9.8 (Molecular Simulations Inc.) with the consistent valence forcefield. In Fig 2(A) peptides SAF-p 1 and SAF-p2 (each coloured dark grey-to-light grey from the N-terminus) interact through core residues including asparagine pairs (coloured mid-grey) to form the two strands of a staggered, parallel, coiled-coil fibre. In Fig. 2(B), negatively charged glutamate side chains (coloured light grey) and positively charged lysine side chains (coloured black) form complementary charge interactions between the SAF peptides.

Enter the attached sequence listing at the end of the specification.

REMARKS

The amendments are made to refer to the sequence identifiers in the sequence listing. No new matter has been entered. Entry is requested.

Respectfully submitted,



Dated: July 22, 2002

Sarah A. Kagan
Registration No. 32,141

BANNER & WITCOFF, LTD.
1001 G Street, N.W.
Eleventh Floor
Washington, D.C. 20001
TEL: (202) 508-9100
FAX: (202) 508-9299

Appendix to show marked-up version of changes made

IN THE SPECIFICATION

At page 4, paragraph 3, substitute the following paragraph:

In a preferred protein structure, the first and second peptide monomer units have the following sequences:

a) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C; SEQ ID NO: 1) and
b) KIRALKAKNAHLKQEIAALEQEIAALEQ (SAF-p2D; SEQ ID NO: 2) respectively; or
c) KIAALKQKIAALKQEIDALEYENDALEQ (SAF-p1A; SEQ ID NO: 3) and
d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 4) respectively; or
e) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C; SEQ ID NO: 1) and
f) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 4) respectively.

At page 9, table legend, substitute the following legend:

*=Chemical capping = CH_3CO at the N terminus and NH_2 at the C terminus. The sequences in the table are SEQ ID NO: 5, 6, 7, 8, 9, 10, 10, 11, and 12, respectively.

At page 7, paragraph 5, substitute the following paragraph:

Fig. 1A and Fig. 1B [illustrates] illustrate the design (Fig. 1A) and the sequences (Fig. 1B; SEQ ID NOs: 15 and 16) of self-assembling fibre (SAF) peptide monomers of the invention.

At page 8, paragraph 2, substitute the following paragraph:

Fig. 8 shows amino acid sequences (SEQ ID NOs: 17 and 18) designed to form blunt-ended heterodimers.

At page 12, paragraph 1, substitute the following paragraph:

In addition and as a control, the SAF-p I c sequence was permuted (N- and C-terminal halves were swapped) to produce peptide SAF-p3:

E IDALEYE NDALEQK IAALKQK IASLKQ (SEQ ID NO: 13.)

This design should combine with SAF-p2D to form a blunt-ended structure, which should not form fibres.

At page 12, paragraph 2, substitute the following paragraph:

A model of the three-dimensional structure of the designed protein fibre resulting from the assembly of SAF-pl and SAF-p2 was made from the minimised structure of a model coiled-coil 35-mer, (LAALAAA)s (SEQ ID NO: 14), which was generated using Crick's Equation and had an ideally packed interface (G. Offer and R. Sessions, J. Mol. Biol. 249, 967 (1995)). Copies of the 35-mer were superimposed with an overlap of one heptad repeat to extend the structural template, and the backbone was rejoined after removal of overlapping°segments. Residues in the two-stranded template were replaced with the sequences of the SAF peptides, staggered relative to each other by two heptad repeats according to the alignment in. Fig. 1B. The structure was soaked in a 5 A layer of water and energy minimised until the average absolute derivative of coordinates with respect to energy fell below 0.01 kcal A-'. The structure was built and visualized using insight II 97.0 (Molecular Simulations Inc.), and was energy-minimized using Discover 2.9.8 (Molecular Simulations Inc.) with the consistent valence forcefield. In Fig 2(A) peptides SAF-p 1 and SAF-p2 (each coloured dark grey-to-light grey from the N-terminus) interact through core residues including asparagine pairs (coloured mid-grey) to form the two

strands of a staggered, parallel, coiled-coil fibre. In Fig. 2(B), negatively charged glutamate side chains (coloured light grey) and positively charged lysine side chains (coloured black) form complementary charge interactions between the SAF peptides.

IN THE CLAIMS:

16. (Amended) A protein structure according to any preceding claim in which the first and second peptide monomer units have the sequence:

a) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-pl; SEQ ID NO: 1) and

b) KIRALKAKNAHLKQEIAALEQEIAALEQ (SAF-p2; SEQ ID NO: 2) respectively; or

c) KIAALKQKIAALKQEIDALEYENDALEQ (SAF-plA; SEQ ID NO: 3) and

d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 4) respectively; or

e) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-pIC; SEQ ID NO: 1) and

f) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 4) respectively.

17. (Amended) A peptide monomer unit for use in preparing a protein structure the peptide monomer unit having an amino acid sequence selected from:

a) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-pl; SEQ ID NO: 1);

b) KIRALKAKNAHLKQEIAALEQEIAALEQ (SAF-p2; SEQ ID NO: 2);

c) KIAALKQKIAALKQEIDALEYENDALEQ (SAF-p1A; SEQ ID NO: 3);

d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 5);

e) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C; SEQ ID NO: 1) ; and

d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C; SEQ ID NO: 5).

PROTEIN STRUCTURES AND PROTEIN FIBRES

This invention relates to protein structures, to methods of producing those protein structures, and to protein fibres and other materials and assemblies produced using those protein structures.

The process of molecular self-assembly is central to all biological systems and is assuming increasing importance and application in biotechnology (L. Q. Gu, *et al* (1999) *Nature* **398**, 686) and nanotechnology (K. E. Drexler, (1999) TIBTECH **17**, 5). The characterization of natural biomolecular assemblies motivates and directs the development of model self-assembling systems and, in turn, these advance our understanding of biology. For proteins at least, the coiled coil is arguably the simplest self-assembling system. Coiled coils are protein-folding motifs that direct and cement a wide variety of protein-protein interactions (A. Lupas, (1996) *Trends Biochem. Sci* **21**, 375). In structural terms, coiled coils are relatively straightforward: they are α-helical bundles with between 2 and 5 strands that can be arranged in parallel, antiparallel or mixed topologies. The basic sequence features that guide the formation of coiled coils from peptides are reasonably well understood (P. B. Harbury *et al* (1993) *Science* **262**, 1401; D. N. Woolfson and T. Alber (1995) *Protein Sci.* **4**, 1596). For instance, most coiled-coil sequences are dominated by a 7-residue repeat of hydrophobic (H) and polar (P) residues, $(HPPHPPP)_n$, known as the "heptad repeat". When configured into an α-helix this pattern gives an amphipathic structure, the hydrophobic face of which directs oligomer-assembly. Furthermore, both the number and the direction of chains within a coiled-coil bundle is determined predominantly by residues that form or flank the hydrophobic core namely, residues at the first, fourth, fifth and seventh positions of the heptad repeat. For instance, coiled coils which form dimers (i.e. two-stranded assemblies) usually have isoleucine or valine residues at the first position and a leucine residue at the fourth position. By contrast, coiled coils that form trimers (i.e. three-stranded assemblies) often have the same residues (i.e both isoleucine or both leucine) at both "H" positions. Finally, hetero-oligomers (that is coiled coils made from strands with different amino-acid sequences) may be directed by complementary charged interactions that flank the hydrophobic core. For these reasons, there have been a number of successful *de novo* protein designs based on the coiled coil These include some ambitious structures that extend the natural repertoire of coiled-coil motifs (S. Nautiyal *et al* (1995) *Biochemistry* **34**, 11645; A. Lombardi *et al* (1996)

Biopolymers **40**, 495; D. H. Lee *et al* (1996) *Nature* **382**, 525; P. B. Harbury *et al* (1998) *Science* **282**, 1462; J. P. Schneider *et al* (1998) *Folding Des.* **3**, R29).

In addition to commonly accepted structures with a single, contiguous heptad repeat, the inventors have identified sequences with multiple, offset heptad repeats which help explain oligomer-state specification in coiled coils. For example, sequences with two heptad repeats offset by two residues; i.e *a/f-b/g'-c/a'-d/b'-e/c'-f/d'-g/e'* set up two hydrophobic seams on opposite sides of the helix formed. Such helices may combine to bury these hydrophobic surfaces in two different ways and form two distinct structures: open "α-sheets" and closed "α-cylinders".

Other relevant aspects of coiled-coil structure are described in WO99/11774, the disclosure of which is incorporated herein by way of reference.

This understanding of coiled coils, and the resulting protein designs, centres on short structures as exemplified by the leucine-zipper motifs (E. K. O'Shea *et al* (1989) *Science* **243**, 538; E. K. O'Shea *et al* (1991) *Science* **254**, 539), which are found in a variety of transcription factors. In contrast, most natural coiled coils extend over hundreds of amino acids (A. Lupas (1996) *supra*; J. Sodek *et al* (1972) *Proc. Natl. Acad. Sci. U.S.A* **69**, 3800) and many assemble further to form thicker, multi-stranded filaments (H. Herrmann and U. Aebi (1998) *Curr. Opin. Struct. Biol.* **8**, 177).

With the goal of making elongated structures to improve our understanding of coiled coils, and to develop protein-design studies, we initially designed two 28-residue peptides — dubbed Self-Assembling Fibre peptides, SAF-p1 and SAF-p2 — to fold and form extended fibres when mixed. Focusing on the buried, hydrophobic-core positions of the structure, rules were incorporated to direct parallel dimer formation and to guard against alternative oligomers and topologies (P. B. Harbury *et al* (1993) *supra*; D. N. Woolfson and T. Alber (1995) *supra*; L. J. Gonzalez *et al* (1996) *Nature Struct. Biol.* **3**, 1011). The building block of the design was a staggered heterodimer with overhanging or "sticky" ends. This contrasts with and distinguishes it from the natural and designer coiled-coil assemblies that have been characterized to date, in which the polypeptide strands align in-register, i.e they have blunt or "flush" ends. Complementary core interactions and flanking ion-pairs were incorporated into the overhangs to facilitate longitudinal association of the heterodimers (Figs. 1&2). This

principle of using "sticky ends" is well developed in molecular biology for assembling DNA (S. J. Palmer *et al* (1998) *Nucleic Acids Res.* 26, 2560), and has been used to design intricate DNA crystals (E. Winfree *et al* (1998) *Nature* 394, 539). However, to our knowledge, our application of sticky end-directed molecular assembly to peptides is new; although we do note that head-to-tail packing of helices has been observed in recently solved crystal structures for two designer peptides (N. L. Ogihara *et al* (1997) *Protein Sci.* 6, 80; G. G. Prive *et al* (1999) *Protein Sci.* 8, 1400). These were helical peptides that crystallised with their helical ends in contact so as to form pseudo-continuous helices in the solid state. In other words they formed "blunt-ended" arrangements.

US-A-5,712, 366 discloses self-assembling protein material but does not provide details of how to make a staggered parallel heterodimer. WO 96/11947 discloses protein nanostructures based on bacteriophage T4 tail fiber proteins but does not disclose a staggered parallel heterodimer coiled coil structure.

Pandya et al., Biochemistry, 29, 8728-34, 2000 (published after the priority date of the present application) does not disclose a method of making nanotubes and does not disclose a matrix comprising the protein structures of the present invention.

According to one aspect of the invention there is provided a protein structure comprising a plurality of first peptide monomer units arranged in a first strand and a plurality of second peptide monomer units arranged in a second strand, the strands preferably forming a coiled-coil structure, and in which a first peptide monomer unit in the first strand extends beyond a corresponding second peptide monomer unit in the second strand in the direction of the strands. The protein structures of the invention have numerous advantages. For example, relatively long protein fibres can be formed with little material - 1 µl of a 100 µM solution of the peptide monomers may provide enough material to form 10 m of fibre 50 nm thick.

At least one charged amino acid residue of the first peptide monomer unit may be arranged to attract an oppositely-charged amino acid residue of the second peptide monomer unit. Preferably, the charged amino acid residue is in an end portion of the first peptide monomer unit, which extends beyond the corresponding second peptide monomer unit in the second strand. At least one strand may consist solely of first or second peptide monomer units

respectively i.e homogenous strands. Heterologous strands are also contemplated. The peptide monomer units may comprise a repeating structural unit. Preferably, the repeating structural unit comprises a heptad repeat motif, having the pattern:

```
hpphppp
abcdefg
```

Preferably, the repeat may include isoleucine or asparagine at position a and leucine at position d. Other repeats (e.g hendecads - abcdefghijk) and amino acid compositions may also be used (see WO99/11774).

Preferably, the heptad repeat comprises oppositely-charged residues at positions e and g respectively. The oppositely-charged residues may be, for example, glutamic acid and lysine residues or arginine and aspartic acid. The use of synthetic amino acids, such as ornithine is also envisaged.

A protein structure in accordance with the invention may be also specified by pairs of asparagine residues in the "a" positions provided by corresponding first and second peptide monomer units.

In a preferred protein structure, the first and second peptide monomer units have the following sequences:

a) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C) and

b) KIRALKAKNAHLKQEIAALEQEIAALEQ (SAF-p2D) respectively; or

c) KIAALKQKIAALKQEIDALEYENDALEQ (SAF-p1A) and

d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C) respectively; or

e) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C) and

f) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C) respectively.

It will be appreciated that these are examples only of 4-heptad structures and that other lengths are possible and envisaged for use in the invention.

According to another aspect of the invention, there is provided a method of producing protein structures, the method comprising providing a mixture of first and second peptide monomer units which associate to form a protein structure according to the invention. The structure can be derivatised and/or stabilized by cross-linking.

Derivatization of the peptide monomer units before or after assembly into the protein structures of the invention may be performed. For example, fluorescent moieties (fluorophores) may be attached to the coiled coil as described in WO99/11774. The addition of fluorescent moieties may assist visualization of the protein structure. Substitution with

unctional groups at the "f" position in the heptad repeat is especially preferred as that position is on the outside of the helix (see Fig. 1C and 1E). Other derivatives may include attaching binders to the peptide monomer units for example so that units which can bind other entities can be produced.

The first and second peptide monomers and the strands may have the characteristics described above.

The invention also provides protein fibres produced by an association of protein structures according to the invention.

The protein structures may also be arranged to form tubular structures. In particular, the structures may be arranged to form nanotubes.

According to another aspect of the invention, there is provided a kit for making protein structures, the kit comprising first and second peptide monomer units which associate to form a protein structure or protein fibres according to the invention.

The protein structures of the invention may be assembled in two and three dimensional arrays. For example, two dimensional mats can be formed which can function, for example as filters. Three dimensional grids or matrices can also be formed again, for example, for use as sieves or filters or for organising other associated or conjugated molecules in three dimensions.

In a preferred embodiment, a matrix is assembled *in situ*. For example, a matrix can be formed in a solution to entrap contaminants in the solution and then the matrix, together with contaminants, can be removed from the solution for example by centrifugation.

The stability of the protein structures at higher temperatures may be improved by making the peptide monomers longer, such that the overlap between corresponding first and second monomer unit residues is increased. Increases in monomer length have previously been shown to stabilize coiled coil structures. Alternatively, stability can be improved by introducing bonding between adjacent peptide monomer units in the same strand. For example, Kent (Dawson *et al* (1994) Science 266: 776) and co-workers have produced peptide bonds between adjacent polypeptide units by coupling and subsequent rearrangement of a

cysteine residue at the N end of one polypeptide unit to a thio-ester derivatised C-terminus of another unit.

Additionally, the protein structures may be stabilised and derivatised by using them to template the polymerisation of synthetic polymers.

Definitions

The terms used in the specification are to be given the ordinary meaning attributed to them by the skilled addressee. The following is given by way of clarification:

Amino acid.

This term embraces both naturally-occuring amino acids and synthetic amino acids as well as naturally-occuring amino acids which have been modified in some way to alter certain properties such as charge. In all cases references to naturally-occurring amino acids may be considered to include synthetic amino acids which may be substituted therefor.

Coiled Coil

A coiled-coil is a peptide/protein sequence usually with a contiguous pattern of hydrophobic residues spaced 3 and 4 residues apart, which assembles (folds) to form a multi-meric bundle of helices. Coiled-coils including sequences with multiple offset repeats are also contemplated.

Dimer

A dimer is a two stranded structure.

Heterodimer

A heterodimer is a dimeric structure formed by two different stands.

Staggered heterodimer

A staggered heterodimer is a structure in which the two strands assemble to leave overlapping ends that are not interacting within the heterodimer.

Blunt-end assembly

Blunt-end assembly is association where the two strands combine to give flushed i.e non-overlapping ends.

Protofibril

A protofibril is a protein structure assembled longitudinally from staggered heterodimers interacting through their overhanging ends.

Fibre

A fibre is a structure formed by lateral association of two or more protofibrils.

Protein structures and methods of producing protein structures in accordance with the invention will now be described, by way of example only, with reference to the accompanying Figures 1 to 8 in which:

Fig. 1 illustrates the design and the sequences of self-assembling fibre (SAF) peptide monomers of the invention.

Fig. 2 illustrates computer modelling of the designed self-assembling fibre of the invention.

Fig. 3 illustrates the results of circular dichroism (CD) and linear dichroism (LD) experiments on protein structures of the invention.

Fig. 4 illustrates the assembly of synthetic protein fibres visualized directly by transmission electron microscopy and an analysis of fibre width In all panels, the white scale bars represent 100 nm. Fig. 4D is a histogram showing the distribution of fibre widths determined using TEM for fresh (white bars) and matured (black bars) mixtures of SAF peptides at 100 μM (a width value of "*x*" on the histogram includes all measurements made from "(*x*–*5*) to *x*").

Fig. 5 is a cartoon showing the possible anti-typic association of parallel helical peptides leading to a homo-oligomeric peptide nanotube.

Fig. 6 is an x-ray diffraction pattern of an aligned protein fibre of the invention.

peptides at 100 μM (a width value of "*x*" on the histogram includes all measurements made from "(*x*–5) to *x*").

Fig. 5 is a cartoon showing the possible anti-typic association of parallel helical peptides leading to a homo-oligomeric peptide nanotube.

Fig. 6 is an x-ray diffraction pattern of an aligned protein fibre of the invention.

Fig. 7 is an image from a confocal fluorescent microscope showing fibres which have been derivatised through the inclusion of flurophores; and

Fig. 8 shows amino acid sequences designed to form blunt-ended heterodimers.

1) Peptide Design and Synthesis

Various peptide monomer units were designed as described above. The monomers and capping peptides (designed to complement the sticky ends of the monomers so as to produce flush, or blunt ends and, so, arrest longitudinal fibre assembly) are set out in Table 1:

TABLE 1

PEPTIDE	SEQUENCE g abcdefg abcdefg abcdefg abcdef 5 10 15 20 25	DESIGN	CD DATA		LD DATA		EM	
			@ 10 μM	@ 100 μM	@ 10 μM	@ 100 μM	@ 10 μM	@ 100 μM
CAP-p1A	*YGPGE IAALEQE NAALEQ	prototype	unfolded					
SAF-p1A	K IAALKQK IAALKQE IDALEYE NDALEQ	prototype; slowly precipitates	unfolded	~ 45 % α-helix				no fibres
SAF-p1B	*K IAALKQK IAALKQE IDALEYE NDALEQ*	chemical capping of the ends (↑ stability)	~ 60 % α-helix	~ 70 % α-helix				
SAF-p1C	K IAALKQK IASLKQE IDALEYE NDALEQ	no capping (↓ stability), mutate $A_{11} \rightarrow S$ (↑ solubility & ↓ helix stability)	unfolded	~ 20 % α-helix		no signal		no fibres
CAP-p2A	K IAALKQK NAALKQG GW*	prototype	unfolded					
SAF-p2A	K ISALKWK NASLKQE IAALEQE IAALEQ	prototype, low solubility	unfolded					
SAF-p2B	*K IRALKWK NAHLKQE IAALEQE IAALEQ*	mutate $S_3 \rightarrow R$ & $S_{11} \rightarrow H$ (↑ solubility & ↑ helix stability)	~ 60 % α-helix	~ 95 % α-helix				
SAF-p2C	K IRALKWK NAHLKQE IAALEQE IAALEQ	no capping (↓ stability)	unfolded	~ 20 % α-helix				thin fibres
SAF-p2D	K IRALKAK NAHLKQE IAALEQE IAALEQ	mutate $W_7 \rightarrow A$ (investigate role of Trp in fibrillogenesis)	~ 15 % α-helix	~ 45 % α-helix		no signal		no fibres
SAF-p2E	K IRALKCK NAHLKQE IAALEQE IAALEQ	mutate $A_7 \rightarrow C$ (for derivatization & cross-linking)						

*=Chemical capping = CH_3CO at the N terminus and NH_2 at the C terminus

SUBSTITUTE SHEET (RULE 26)

Fig. 1 shows (A) A mechanism for self-assembly: complementary charges in "companio peptides direct the formation of staggered, parallel heterodimers; the resulting "sticky" en are also complementary and promote longitudinal association into extended structures. F 1(B) shows the designed amino acid sequences: each peptide comprised canonical hept repeats (***abcdefg***) with Ile at ***a*** and Leu at ***d*** to guide the formation of coiled-coil dime oppositely-charged residues were incorporated at ***e*** and ***g*** to favour the staggered dimer wi sticky ends; asparagine residues (which preferentially pairs with each other at ***a*** sit (Gonzalez L *et al* (1996) Nature Structural Biology 3, **13**: 1011-1018) were included cement the prescribed register further and to favour the parallel structures. Fig. 1(C) is helical-wheel representation, summarizing the designed sequences in context. The view from the N-terminus with heptad sites labeled ***a-g*** and assumes 3.5 residues per helical turn t emphasise the heptad repeat.

The peptides were synthesized on an Applied Biosystems 432A Peptide Synthesizer usin solid-phase methods and Fmoc chemistry. Peptide samples were purified usin reversed-phase HPLC and their identities confirmed by MALDI-TOF mass spectrometry.

Various combinations of peptide monomers and capping peptides were tested as set out i Table 2:

TABLE 2

MISC.	PEPTIDE MIXTURE	EQUILIBRATION	CD DATA @ 10 μM	CD DATA @ 100 μM	LD DATA @ 10 μM	LD DATA @ 100 m μM	EM DATA @ 10 μM	EM DATA @ 100 μM
	SAF-p1A, SAF-p2A		~ 20 % α-helix	~ 40 % α-helix $T_m \approx 30$ °C				
	SAF-p1A, SAF-p2A, CAP-p1A, CAP-p2A			~ 20 % α-helix				
	SAF-p1A, SAF-p2B	rapid		~ 65 % α-helix $T_m \approx 36/46$ °C (expect 38 °C if no interaction)				
	SAF-p1B, SAF-p2B	rapid	~ 70 % α-helix $T_m \approx 25$ °C (expect 19 °C if no interaction)	~ 80 % α-helix $T_m \approx 35$ °C (expect 36 °C if no interaction)				no fibres
	SAF-p1A, SAF-p2C	slow @ 100 μM; clouding occurs	~ 20 % α-helix	~ 80 % α-helix biphasic thermal melt			no fibres	thick fibres (~ 45 nm wide)
no signal from fibres in 1D-NMR	SAF-p1C, SAF-p2C	slow @ 100 μM; clouding occurs	~ 25 % α-helix	~ 65 % α-helix (~25% in 0 5 M salt), unusual spectral shape; no aromatic signal	no signal	strong signal from backbone & aromatics	no fibres	thick fibres (~ 45 nm wide); no fibres in 0 5 M salt
	SAF-p1C, SAF-p2D	slow @ 100 μM; clouding occurs	~ 25 % α-helix	~ 85 % α-helix (~20% in 0 5 M salt), unusual spectral shape	no signal	strong signal from backbone & aromatics, no signal in 0 5 M salt	no fibres	thick fibres (~ 45 nm wide); no fibres in 0.5 M salt

Indicates large structures are formed.

In addition and as a control, the SAF-p1c sequence was permuted (N- and C-terminal halves were swapped) to produce peptide SAF-p3:

E IDALEYE NDALEQK IAALKQK IASLKQ

This design should combine with SAF-p2D to form a blunt-ended structure, which should not form fibres.

2) Modeling of Protein Fibre Structure

A model of the three-dimensional structure of the designed protein fibre resulting from the assembly of SAF-p1 and SAF-p2 was made from the minimised structure of a model coiled-coil 35-mer, $(LAALAAA)_5$, which was generated using Crick's Equation and had an ideally packed interface (G. Offer and R. Sessions, *J. Mol. Biol.* **249**, 967 (1995)). Copies of the 35-mer were superimposed with an overlap of one heptad repeat to extend the structural template, and the backbone was rejoined after removal of overlapping segments. Residues in the two-stranded template were replaced with the sequences of the SAF peptides, staggered relative to each other by two heptad repeats according to the alignment in Fig. 1B. The structure was soaked in a 5 Å layer of water and energy minimised until the average absolute derivative of coordinates with respect to energy fell below 0.01 kcal $Å^{-1}$. The structure was built and visualized using Insight II 97.0 (Molecular Simulations Inc.), and was energy-minimized using Discover 2.9.8 (Molecular Simulations Inc.) with the consistent valence forcefield. In Fig 2(A) peptides SAF-p1 and SAF-p2 (each coloured dark grey-to-light grey from the N-terminus) interact through core residues including asparagine pairs (coloured mid-grey) to form the two strands of a staggered, parallel, coiled-coil fibre. In Fig. 2(B), negatively charged glutamate side chains (coloured light grey) and positively charged lysine side chains (coloured black) form complementary charge interactions between the SAF peptides.

3) Circular Dichroism Experiments

Peptide samples were incubated at 5°C in 10 mM MOPS (3-(N-Morpholino)propanesulfonic acid), pH 7. Sample concentrations were determined from their UV absorbance at 280 nm (SAF-p1) and 214 nm (SAF-p2). After baseline correction, ellipticities in mdeg were converted to molar ellipticities (deg cm^2 dmol-res^{-1}) by normalizing for the concentration of

peptide bonds. Data were recorded in a cell of 1 mm path length by integrating the signal for 5s (and 1s for the fresh 100 μM peptide mixture) every nm in the range 205-260 nm. CD measurements were made using a JASCO J-715 spectropolarimeter fitted with a Peltier temperature controller.

The CD data shown in Fig. 3 provides spectroscopic evidence for the formation of helical structures by the SAF peptides. Fig. 3(A) shows circular dichroism (CD) spectra at 10 μM for: SAF-p1 (----), SAF-p2 (- -), the average of these spectra (- - -), and the experimental SAF peptide mixture (o). Fig 3(B) shows CD spectra at 100 μM - the key is the same as for Fig 3(A), but with the additional spectrum (•) being for the SAF peptide mixture after "maturation" for 1 h.

Consistent with our design, neither SAF-p1 nor SAF-p2 was highly structured in aqueous solution at pH 7 and 5 °C (Fig. 3). However, when mixed in equal proportions the circular dichroism (CD) spectrum changed and, moreover, was markedly different from the theoretical spectrum generated by averaging the spectra for the isolated peptides. In particular, the spectrum for the mixture had intense minima at 208 and 222 nm consistent with the formation of α-helical structure, but these features were not as pronounced in the spectra of the individual peptides. This was clear evidence that the two peptides interacted to form an α-helical structure as designed. Furthermore, and as expected for a multimerization event, the magnitude of these spectral changes depended on peptide concentration; a SAF mixture with 10 μM of each peptide, did show a weak signal indicative of some α-helical structure, however, a 100 μM mixture gave a much stronger signal (Figs. 3A&B).

The shape and intensity of spectra from 100 μM mixtures of the SAF peptides also changed with time (Fig. 3B). Spectra recorded immediately after mixing a "fresh" sample displayed some α-helical structure. After incubation of the mixture for 1 hour at 5 °C ("maturation"), however, the signal at 222 nm was more intense, and indicated approximately 75 % α-helix, consistent with substantial coiled-coil formation.

Maturation of 100 μM SAF peptide mixtures was also accompanied by slight clouding of the samples. Scattering effects from such samples can lead to attenuation and distortion of CD spectra (D. Mao and B. A. Wallace, (1984) *Biochemistry* **23**, 2667). However, we could

disregard this possibility because altering the distance between the sample and the detector in the CD instrument did not affect the shape or the intensity of the spectrum. Furthermore, we established that the majority of the CD signal from the mixtures derived from the suspended material: a supernatant without the suspended material, which was recovered by centrifugation of a matured 100 μM SAF mixture, gave only a weak CD signal similar to the 10 μM mixture.

Thus, the CD data were wholly consistent with the desired α-helical SAF design and, moreover, indicated the formation of large assemblies.

As a control, SAF-p3 (the permutation of SAF-p1 (identical to SAF-p1c)) was designed to form a blunt-ended heterodimer with SAF-p1 that should not assemble further into fibres. 100 μM mixtures of SAF-p2 (identical to SAF-p2D) and SAF-p3 were analysed by sedimentation equilibrium in the analytical ultracentrifuge. The resulting data were best fitted assuming a single ideal species in solution, and the molecular weight was allowed to vary during the fit. An M_r of 6422 (with 95% confidence limits of 5924 and 6911) was obtained, which is very close to the expected heterodimer value of 6303 calculated from mass spectrometry of the individual peptides. CD spectra for 100 μM fibre-producing mixtures (SAF-p1 with SAF-p2), and for blunt dimer-producing mixtures (SAF-p2 with SAF-p3), were recorded. For the blunt dimer-producing mixtures, the shape and intensity of the CD spectrum were fully consistent with coiled-coil formation as designed. In contrast to the fibre-producing mixtures, the blunt dimer-producing mixtures showed no signs of maturation; that is, negligible spectral changes and no clouding of solutions occurred upon incubation. Interestingly, the intensity of the minimum near 222 nm, which is an accepted indicator of α-helical structure and degree of α-helical folding, was similar for both mixtures. This strongly supports the formation of α-helical structure as designed in the fibre-producing mixtures despite the spectral shifts observed upon maturation.

4) Linear Dichroism Experiments

Linear dichroism (LD) spectroscopy was also used to test if elongated structures were being formed as designed. Long polymers such as DNA molecules can be oriented by shear flow. This effect can be monitored by LD spectroscopy provided that chromophores also become

aligned by the flow (M. Bloemendal (1994) *Chem. Soc. Rev.* **23**, 265; A. Rodger and B. Norden (1997) Oxford Chemistry Masters (Oxford University Press, Oxford), vol. 1).

Peptide samples were prepared for LD as for CD. LD data were collected on samples spinning in a couette flow cell by integrating the signal for 2 s every nm in the range 210-320 nm, using a JASCO J-715 spectropolarimeter. After baseline correction, absorbance was converted to molar extinction coefficient (1 mol-res^{-1} cm^{-1}) by normalizing for the concentration of peptide bonds. A linear correction for a sloping baseline was made to the data from the 100 μM SAF peptide mixture.

The results are depicted in Fig. 3D, which shows linear dichroism (LD) spectra for: 20 μM tropomyosin (- - -), the SAF peptide mixture at 10 μM (----), and the SAF peptide mixture at 100 μM in the absence (•) and presence (o) of 0.5 M KF.

For instance, we found that tropomyosin, which forms a dimeric coiled coil approximately 42 nm in length, could be aligned to give a LD signal (Fig. 3D). In contrast and consistent with our design and the CD data, experiments with a 10 μM SAF mixture, (Fig. 3D), and for the individual peptides at 100 μM (data not shown), LD signals were not detected. However, a matured 100 μM SAF peptide mixture gave a strong absorbance from the peptide backbone (210-240 nm) and some signal in the aromatic region (260-290 nm) during flow orientation (Fig. 3D). As only long structures are aligned by this technique, the data demonstrated that long fibres at least 500 nm in length were present in solutions of the matured 100 μM SAF peptide mixtures.

5) Electron Microscopy

To confirm fibre assembly, we used electron microscopy to visualize structures in the peptide preparations directly. For TEM experiments, peptide samples were incubated for 1 h at 5 °C in filtered 10 mM MOPS, pH 7. A drop of peptide solution was applied to a carbon-coated copper specimen grid (Agar Scientific Ltd, Stansted, UK), and dried with filter paper before negative staining with 0.5% aqueous uranyl acetate and then dried at 5 °C. A "fresh" SAF peptide mixture was prepared by mixing preincubated solutions of the individual peptides at 200 μM directly on the specimen grid, before drying and negative staining as described. Grids were examined in a Hitachi 7100 TEM at 100 kV and digital images were acquired

with a (800 x 1200 pixel) charge-coupled device camera (Digital Pixel Co. Ltd., Brighton, UK) and analyzed (Kinetic Imaging Ltd., Liverpool, UK).

For scanning electron microscopy (SEM) experiments, negatively-stained specimen grids were sputter-coated with gold and examined in a Leo Stereoscan 420 SEM at 20 kV and with a probe current of 10 pA.

No structures were visible up to 100 000 times magnification by transmission electron microscopy (TEM) for either the 10 μM SAF mixture, or for the individual peptides at 100 μM concentration (data not shown). However, TEM of a 100 μM SAF mixture at 50 000 times magnification revealed time-dependent formation of long fibrous structures, consistent with the CD and LD data. Fresh mixtures showed large numbers of extended fibres of various widths. The majority of these had a diameter of about 20 nm (Figs. 4A (a fresh mixture at 100 μM) & Fig 4D); finer fibres were present, but their widths could not be measured reliably. Images recorded for the matured mixtures showed fewer fibres, but these were more distinct and thicker than those observed in the fresh mixture (Fig. 4B&D).

Scanning electron microscopy (SEM) of a matured mixture showed no evidence for fibre branching. Rather, the fibres were simply intertwined as if layered on top of each other (Fig. 4C). It was not possible to follow the full length of fibres due to intertwining, but they were at least several hundred microns in length. Although the density of fibres varied across the surface of the EM grid, for the matured samples at least, their diameters were quite uniform with a mean width of 43.3 (SD = 9.3) nm (Fig. 4D). As the original design was for a longitudinally extended, but otherwise two-stranded coiled coil the average diameter that we might have expected was about 2 nm. Therefore, the EM data suggested that the designed two-stranded coiled-coil fibres associate laterally into higher order assemblies.

6) X-ray Fibre Diffraction

Mixtures of SAF peptides at 500 μM in 10 mM MOPS, pH 7, were incubated on ice for at least 1h, before centrifugation at 6500g for 5 min. Droplets of fibre-containing solutions, taken from the bottom of the centrifuged tubes, were suspended between the ends of two wax-filled capillaries and allowed to dry slowly overnight at 4°C, yielding clumps of partially aligned fibres. X-ray fibre diffraction images were collected using a Rigaku CuKα rotating

anode source (wavelength 1.5418 Å) and a R-AXIS IV detector. Samples were maintained at 5°C during data collection with cool air from a cryostream (Oxford Cryo-systems). The X-ray fibre diffraction pattern collected from SAF peptide fibres showed the following features (Fig. 6): (1) a short meridional (that is, parallel to the long fibre axis) reflection at 5.11± 0.03 Å; (2) the harmonic of this 5.11 reflection at 10.19 ± 0.05 Å; and (3) a stronger, more diffuse reflection centered at 8.8 ± 0.15 Å on the equator. These features are consistent with α-helical coiled-coils aligned with the fibre axis. The 5.1 Å meridional reflection corresponds to the pitch of the helices within the coiled-coils. The other expected reflection on the meridian-that is, that at 1.5 Å and corresponding to the rise per residue-lies out of the resolution of the current data sets, whereas the equatorial reflection reveals the mean distance between α-helical axes. This value at 8.8 Å is less than the observed value for keratin but falls within reported ranges for dimeric coiled-coil peptides.

7) Effect of Potassium Fluoride on Protein Fibre Assembly

Molecular modeling of the SAF sequences into an extended two-stranded coiled coil also highlighted potential complementary charge interactions on the surface of the protofibrils, Figs 1&2. In accordance with this, experimentally it was found that moderate concentrations of salt inhibited protofibril and thick fibre assembly. First, CD spectra recorded for both the individual peptides and a 100 μM mixture of SAF peptide samples with 0.5 M potassium fluoride showed reduced helical CD signals and there was no evidence of "maturing" in the mixed samples (Fig. 3C). Second, the LD signal described previously for the matured 100 μM SAF peptide mixture was also lost when the experiment was repeated in the presence of salt (Fig. 3D). Finally, TEM images of a 100 μM SAF mixture also demonstrated that fibres were not formed in the presence of 0.5 M KF (Fig. 4E). Fig. 4E shows the results of TEM of a matured SAF peptide mixture at 100 μM incubated in the presence of 0.5 M KF.

The inventors did not knowingly design any features into the SAF peptides to foster further association of the two-stranded coiled coils. The observation of thick fibres in SAF peptide preparations, therefore, raised the question: what interactions guided and stabilized these higher-order assemblies? The inventors therefore propose that features inherent in repeating structures of the type that they designed will naturally promote such fibre assembly (fibrillogenesis).

Consider a protofibril as depicted in Fig. 1B and 2A. Any sequence feature presented on its surface by either, or both of the constituent peptides will be repeated at regular intervals along the protofibril. The repeat length will be equal to the length of the peptides (for SAF-p1 and SAF-p2 this was 28 residues, or about 4.2 nm). Furthermore, the motif will spiral around the protofibril tracking the superhelix of the coiled coil, which has a pitch of about 15 nm for a contiguous, heptad-based, dimeric structure. In this scenario, protofibril-protofibril interactions may be promoted if another sequence motif complementary to the first is present in the potential partner. This is because the pitches of the complementary motifs on each protofibril will match precisely. Thus, once initiated, lateral association of protofibrils — that is, fibrillogenesis — will be cemented by many regularly spaced interactions as in a crystal. As a result, the complementary interactions need only be weak as the stability of the protofibril-protofibril interaction rests on an avidity effect rather than a small number of strong interactions. Provided that the components of the assembly can make more than one type of complementary surface very extensive molecular assemblies may result.

The inventors used electrostatic interactions both to direct heterodimer formation, and to promote elongation of the protofibrils (Figs. 1 and 2). These features would also create periodic and alternating patches of charge in the protofibrils provided they are regular as envisaged (Fig. 1B and 2B). These charged patches could guide and stabilize the higher order assemblies. Indeed, similar features have been noted in several natural fibrous proteins and have been implicated in the assembly of multi-protein filaments (J. J. Meng *et al* (1994) *Biol. Chem.* **269**, 18679; A. D. McLachlan and M. Stewart (1976) *Mol. Biol.* **103**, 271), and small synthetic peptide systems (S. G. Zhang *et al* (1993) *Proc. Natl. Sc. U.S.A* **90**, 3334). The experiments with salt (KF) described above suggest that salt-bridges (electrostatic interaction) may be at least in part the cause of fibrillogenesis.

8) Coiled-coils design

a. For two superimposed heptads there are three possible sequence offsets of 1, 2 and 3 residue(s), which are equivalent to 6, 5 and 4-residue offsets, respectively. For a regular 3.6-residue-per-turn α-helix, these set up two hydrophobic faces with angular offsets of 100°, 160° (360-200) and 60° (360-300), respectively, around the outside of the helix. This is best seen on a helical wheel. Accounting for helical supercoiling - i.e assuming 3.5 residues per turn and using the accepted helical-wheel representation for the

coiled-coil these angular offsets are altered to 103°, 154° and 51°, respectively. However, both sets of angles are over-simplifications when considering helix-helix interactions in actual coiled-coil systems because side-chain size, geometry and packing also affect the helix interfaces (Harbury, P. B. *et al* (1993) Science **262**, 1401-1407; Harbury, P. B. *et al* (1994) Nature **371**, 80-83; Malashkevich, V. N. *et al* (1996) Science **274**, 761-765). Nonetheless, we found that many natural coiled-coil assemblies, at least, were consistent with the approximate angular offsets: Trimers could be considered as having overlapping heptads separated by 3 residues (angular offset = 51/60°). Whereas, tetrameric and pentameric coiled-coils were often variations on a theme with two hepad repeats offset by 1 residue (100/103°).

b. **Two heptad repeats offset by two residues: α-cylinder constructions**

Sequence offsets of 2 residues are potentially more interesting than the 1- and 3-residue offsets. This is because of the possibility of placing hydrophobic (H) residues at a, c, d, and f, with c and f effectively making up the a' and d' positions of the second, offset heptad. This is represented below, where P signifies polar (non-core) residues.

a b c d e f g a b c d e f g	repeat 1
H P P H P P P H P P H P P P	binary pattern 1
P P H P P H P P P H P P H P	binary pattern 2
f'g'a'b'c'd'e'f'g'a'b'c'd'e'	repeat 2
a b c d e f g a b c d e f g	assigned register
H P H H P H P H P H H P H P	overall binary pattern

Such sequence patterns would results in two hydrophobic seams with a wide angular separation (154/160°), which would place them roughly on opposite sides of the helix. Furthermore, it offers two possibilities for parallel helix-helix packing arrangements: *syn*, where two like faces - i.e a / d with a / d , or c/f with c/f - from neighbouring helices combine to produce an openα-sheet, Fig. 6a; anti, where a/d faces pair with c/f. In the anti-arrangement the structure can close to form α α-cylinder. For antiparallel pairs of helices syn-typic association should lead to cylinders, whereas sheets should be formed

from anti-typic antiparallel interfaces.

c. **A natural α-cylinder**

TolC has two α-barrel-like domains (Koronakis, V. *et al* (2000) Nature **405**, 914-919). Both have 12 helices contributed by 3 monomers. In the lower barrel each helix pairs with another from the same protomer to form separate supercoiled, antiparallel coiled-coils; SOCKET analysis revealed extensive antiparallel knobs into holes (KIH) interactions within these pairs, but not between them. In contrast, the helices of the upper barrel appear to pack more uniformly, albeit with a slant, to describe an α-cylinder. The SOCKET output for this part of the structure revealed many fewer KIH interactions than found in the lower barrel. Furthermore, KIH interactions were not contiguous around the cylinder and, in particular, they were more extensive between helices in the same monomer, but less regular between the helices abutting the monomers. In our view, the TolC barrel represents a variation of the cylinders formed by protein structures of the invention.

Nevertheless, the inventors were able to assign heptad registers for the helices of the upper barrel unambiguously. This revealed knobs at relative *a*, *c*, *d*, and *f* positions and syn-typic association of two seams adjacent helices; i.e fully consistent with the theory outlined above.

We believe that it will be possible to constuct α-sheets and α-cylinders using helices in parallel. The use of parallel helices does have one interesting consequence for the construction of α-cylinders, however: as the pairing in these structures will be anti-typic, *a* residues on one helix partner *c* residues of a neighbouring helix at the same level in the structure. Similarly, *d* and *f* residues pair at the intervening levels. The result will be that successive helices will be translated up the helix and cylinder axes by two residues, which is equivalent to ≈3Å. Thus, attempts to construct α-cylinders from parallel helices will give spirals of helices which may or may not close. This is, however, potentially extremely interesting as it opens up possibilities for making peptide-based nanotubes as described above.

A second consideration for α-cylinder construction is the consequences of helix and coiled-coil supercoiling. The upper barrel of TolC has 12 helices. Based on a structure of parallel helices with canonical supercoiling, i.e an angular separation of 154° between the two seams in each helix, we calculated that the cylinder should close at 14 helices. However, variations in helix number are expected. One reason for this is that helices cannot supercoil in two direction simultaneously, and some distortion is required to maintain packing at both interfaces. We found structural precedents for this in the Protein Data Book PDB where tight knobs-into-holes packing was maintained (Walshaw & Woolfson, unpublished); indeed, the central helices of the 3-helix α-sheets are straight, Fig. 7b. (n.b . The slanting of the helices in the upper barrel of TolC may offer a compromise between straight and supercoiled helices). Assuming the packing of completely straight helices, the angular offset becomes 160° and 18 helices would close a cylinder. However, given that, as in 3-, 4- and 5-stranded coiled coils, side chains mediate the helix-helix contact angles other oligonmerisation states might be possible (Harbury, P. B *et al* (1993) Science **262**, 1401-1407; Harbury, P. B. *et al* (1994) **371**, 80-83; Malashkevich, V. N. *et al* (1996) Science **274**, 761-765): we calculate that small adjustments in the angular offset between 144° to 162° varies the helix number from 10 to 20.

9) Formation of Protein Structures

As mentioned above, the protein structures of the invention may have various applications such as in:

Nanotubes

a. This can be achieved for example by combining the aforementioned 7- and 11-residue repeats with offsets in the sequence. The effect would be eliminate the overall hydrophobic displacement. In other words, alternating heptad and hendecad repeats give an 18-residue repeat to match the α-helical repeat; in the α-helix, 18 residues span 5 helical turns exactly. It may therefore be possible to create a completely closed peptide nantotube (Fig. 5 shows part of a nanotube) In the parallel, straight helix case there would be 18 helices per turn of the "cylinder", and the rise per turn is 36 residues. Thus, a 36-residue peptide with a 7-11-7-11 repeat offset by 2 residues should form a spiral of

helices the ends of which meet to close the tube. Such nanotubes maybe particularly useful in the production of nanoscale piping and plumbing. The interior of the tube may be derivatised to control the flow of different small (2-40Å) molecules.

b. **Derivitised and branched peptides and peptide templates**

The self-assembling peptides of the invention are relatively small and synthetically accessible. Thus, non-standard derivatisable side chains may be incorporated in them. For example, the monomer units can be made with a single cysteine residue at an exterior f position. These can be used to couple small molecules and other peptides using thiol-based chemistry. A wide variety of thiol-reactive probes are available. In particular, the peptides can be tagged with fluorophores. For instance, with one peptide labelled with Fluorescein and the other with Rhodamine fibres visualised by confocal microscopy appear green and red, respectively (Fig. 7). There is a possibility for FRET between the probes, which may pack closely in the fibres, and this may confuse interpretation. To avoid this the tagged peptides can be doped into fresh, assembling SAF mixtures. Having available fluorescently labelled peptides and fibres offers another route to tracking fibre/network assembly and orientation.

To generate branched self-assembling fibres "T-shaped" conjugated peptides can be made. These are covalent heterodimers made by mixing and coupling together variants of two SAF peptides: one with a terminal cysteine and the other having a central cysteine residue. The desired products can be purified from the mix of disulphide-linked peptide by PHLC. Doping the conjugated ("T") peptides into fresh SAF mixtures should propagate fibre assembly in three dimensions as both the "bar" and the "stem" of the "T" could become incorporated in, or initiate, fibres. The resulting networks can be visualised and characterised by EM.

Peptide synthetic diblock copolymer hybrids may be produced. Suitable methods for preparing water soluble diblock copolymers using atom transfer radical polymerisation are described in X. S. Wang *et al* Chemical Communications 1817 (1999) and X. S Wang *et al* Macromolecules 33, 257 (2000).

The protein fibres of the invention may be used to template and control this

polymerisation either to produce hybrid fibres or if the peptide template is subsequently disassembled and marked away, to provide routes to water soluble "fishnet" nanotubes. Other possibilities include: conjugating polymers onto preassembled peptide fibres; conjugating the polymers and peptides prior to fibre assembly; and effecting polymerisation on the pre-assembled fibres.

c. **Formation of Matrices**

The protein fibres of the invention may be arranged to form two and three dimensional grids and matrices respectively. One application for such matices is in the purification of biological fluids such as blood. An affinity matrix could be assembled (for example *in situ* in blood) to remove blood contaminants such as viruses. In the case of virus removal, a binder for the target contaminant (e.g a peptide or protein with natural or engineered affinities for a viral coat protein) can be fused to a peptide monomer units in the protein structure of the invention. The matrix can then be removed from blood along with any bound contaminants by light centrifugation. For example, it is estimated that a 100 nm length of fibre would have a mass of $\geq$ 12 MDa which would readily be removed. Such affinity matrices have a number of advantages over larger naturally occurring proteins. In the assembled matrices any binders are aligned to give high effective avidities for the targeted molecules.

d. **Other applications**

Other applications for protein structures in accordance with the invention include:

i. preparation of organised networks for seeding the crystalisation of biomolecules for X-ray crystallography;
ii. using ordered fibres to promote cell growth for tissue engineering;
iii. the construction of nanoscale molecular sieves
iv. the preparation of nanoscale molecular grids/scaffolds that could be used as supports for a variety of functional small or macromolecules.
v. functionalised grids and networks could be used in, for example, catalysis, affinity-sieving/purification of biological fluids and other research solutions, the recruitment of endogenous molecules and co-factors to promote tissue repair and tissue engineering in general.

vi. to create novel lab-on-chip technologies, peptide self-assembly could be combined with lithography as follows.

Lithography and related techniques can be used to pattern a variety of surfaces with channels, which can be made of a suitable size (e.g 20-100 nm wide and deep) to accommodate peptide fibres. These can then be used to direct the assembly of the fibres from solutions mixed directly on the surfaces. Furthermore, using well-established chemistry, the inventors envisage funtionalising the peptide fibres with a variety of small molecules and other proteins. This proposed combination of peptide design, self-assembly and lithography should allow the development of ordered arrays of functional polymers on specific surfaces.

vii. Assembled fibres could also be used as fine (therefore, high resolution) tips in AFM (atomic force microscopy) the current limit is about 10-25 nm using carbon nanotubes.

Claims

1. A protein structure comprising a plurality of first peptide monomer units arranged in a first strand and a plurality of second peptide monomer units arranged in a second strand wherein the first and second monomer units comprise the heptad repeat motif (abcdefg) and/or the hendecad repeat motif (abcdefghijk), and wherein a pair of asparagines, arginines, lysines or other complementary residues in the "a" position on at least one pair of corresponding first and second monomer units ensures that the first strand and the second strand form a staggered parallel heterodimer coiled coil structure.

2. A protein structure according to claim 1, wherein a first peptide monomer unit in the first strand extends beyond a corresponding second peptide monomer unit in the second strand in the direction of the strands.

3. A protein structure according to any one of claims 1 to 2 in which at least one charged amino acid residue of a first peptide monomer unit is arranged to attract an oppositely-charged amino acid residue of a second peptide monomer unit.

4. A protein structure according to claim 3 in which the charged amino acid residue is in an end portion of the first peptide monomer unit which extends beyond the corresponding second peptide monomer unit in the second strand.

5. A protein structure according to any one of the preceding claims in which at least one strand consists solely of first or second peptide monomer units respectively.

6. A protein structure according to any one of the preceding claims wherein one or more of the other "a" positions of the first and second monomer units is a hydrophobic residue.

ART 34 AMDT

7. A protein structure according to claim 6, wherein the hydrophobic residue is selected from isoleucine or valine.

8. A protein structure according to any one of the preceding claims having a leucine at one or more of the "d" positions of the first and second monomer units.

9. A protein structure according to any one of the preceding claims having oppositely-charged or otherwise complementary residues at positions g and e of respective monomer units.

10. A protein structure according to claim 9 in which the oppositely-charged residues are glutamic acid and lysine residues or arginine and aspartic acid residues, or synthetic derivatives of these amino acid residues.

11. A protein structure according to any preceding claim in which the structure is stabilised by pairs of asparagine, arginine, lysine or other complementary residues provided by corresponding first and second peptide monomer units.

12. A protein structure according to any preceding claim which is arranged to form a tubular structure.

13. A protein structure according to claim 12 in which the repeat motifs are offset by two or more amino acid positions in sequence whereby the peptide monomer units form a cylinder.

14. A protein structure according to any preceding claim in which the first and second peptide monomer units have the sequence.

a) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1) and

b) KIRALKAKNAHLKQEIAALEQEIAALEQ (SAF-p2) respectively; or

c) KIAALKQKIAALKQEIDALEYENDALEQ (SAF-p1A) and

d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C) respectively; or

e) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C) and

f) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C) respectively.

15. A peptide monomer unit for use in preparing a protein structure the peptide monomer unit having an amino acid sequence selected from:

a) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1);

b) KIRALKAKNAHLKQEIAALEQEIAALEQ (SAF-p2);

c) KIAALKQKIAALKQEIDALEYENDALEQ (SAF-p1A);

d) KIRALKWKNAHLKQEIAALEQEIAALEQ (SAF-p2C) and

e) KIAALKQKIASLKQEIDALEYENDALEQ (SAF-p1C).

16. A protein structure according to any one of claims 1 to 14 or a peptide monomer unit according to claim 15 wherein at least one amino acid residue is derivatised.

17. A branching self-assembling fibre comprising two or more protein structures according to any one of claims 1 to 11, coupled together to form a T-shaped conjugated structure.

18. The branching self-assembling fibre of claim 17, wherein at least one of the protein structures comprises one or more central cysteine residues, and at least one other protein structure comprises a terminal cysteine residue.

19. A method of producing protein structures, the method comprising providing a mixture of first and second monomer units which associate to form a protein structure according to any one of claims 1 to 14, wherein the first and second monomer units comprise the heptad repeat motif (abcdefg) and/or the hendecad repeat motif (abcdefghijk).

20. A method according to claim 19 in which the protein structure is derivatised.

21. A method according to claim 19 or 20 in which the protein structure is stabilised by cross-linking.

22. A protein fibre produced by an association of protein structures according to any one of claims 1 to 14.

23. A kit for making a protein structure, the kit comprising first and second peptide monomer units which associate to form a protein structure according to any one of claims 1 to 14 or a protein fibre according to claim 22, wherein the first and second monomer units comprise the heptad repeat motif (abcdefg) and/or the hendecad repeat motif (abcdefghijk).

24. A two dimensional grid comprising a protein structure according to any one of claims 1 to 14 or a protein fibre according to claim 22.

25. A three dimensional matrix comprising a protein structure according to any one of claims 1 to 14 or a protein fibre according to claim 22.

26. A matrix according to claim 25 which is arranged to assemble in solution.

27. A matrix according to claim 25 or claim 26, wherein one or more binders is fused to the protein structure, wherein the one or more binders are aligned to give high avidities for one or more target entities.

28. A matrix according to any one of claims 25 to 27 which is arranged to bind one or more target entities.

29. A matrix according to claim 28 which is arranged to bind viruses.

30. A method of forming a matrix according to any one of claims 25 to 29 in which a mixture of separate first and second monomer units is provided, wherein the first and second monomer units comprise the heptad repeat motif (abcdefg) and/or the hendecad repeat motif (abcdefghijk) and are caused to associate to form a plurality of protein structures according to any one of claims 1 to 14, wherein the protein structures assemble to form a three-dimensional matrix.

31. A method according to claim 30 in which the matrix is formed *in situ*.

32. A method for controlling the production of a synthetic polymers comprising assembling a protein structure in accordance to any one of claims 1 to 14 in association with the polymer.

33. A method according to claim 32 in which the protein structure is removed after synthesis of the polymer.

34. A tip for use in Atomic Force Microscopy comprising a protein structure according to any one of claims 1 to 14.

Peptides

This invention relates to protein fibre formation and in particular to methods of producing protein fibres to form a protein structure comprising a plurality of first polypeptide units arranged in a first polypeptide strand and a plurality of second polypeptide units arranged in a second polypeptide strand, the strands preferably forming a coiled coil structure, and in which a first polypeptide unit in the first strand extends beyond a corresponding second polypeptide unit in the second strand in the direction of the strands.

(12) INTERNATIONAL [redacted] ICATION PUBLISHED UNDER THE PAT[redacted] COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 March 2001 (29.03.2001)

PCT

(10) International Publication Number
WO 01/21646 A1

(51) International Patent Classification[7]: C07K 14/00, 1/113, G11B 9/00

(21) International Application Number: PCT/GB00/03576

(22) International Filing Date:
18 September 2000 (18.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9922013.9 17 September 1999 (17.09.1999) GB

(71) Applicant *(for all designated States except US)*: UNIVERSITY OF SUSSEX [GB/GB]; Falmer, Brighton, Sussex BN1 9QG (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: WOOLFSON, Derek, N. [GB/GB]; University of Sussex, Falmer, Brighton, Sussex BN1 9QG (GB). WALSHAW, John [GB/GB]; University of Sussex, Falmer, Brighton, Sussex BN1 9QG (GB). PANDYA, Maya, J. [GB/GB]; University of Sussex, Falmer, Brighton, Sussex BN1 9QG (GB). COLYER, John [GB/GB]; Elfordlea, Mill Lane, Bardsey LS17 9AN (GB).

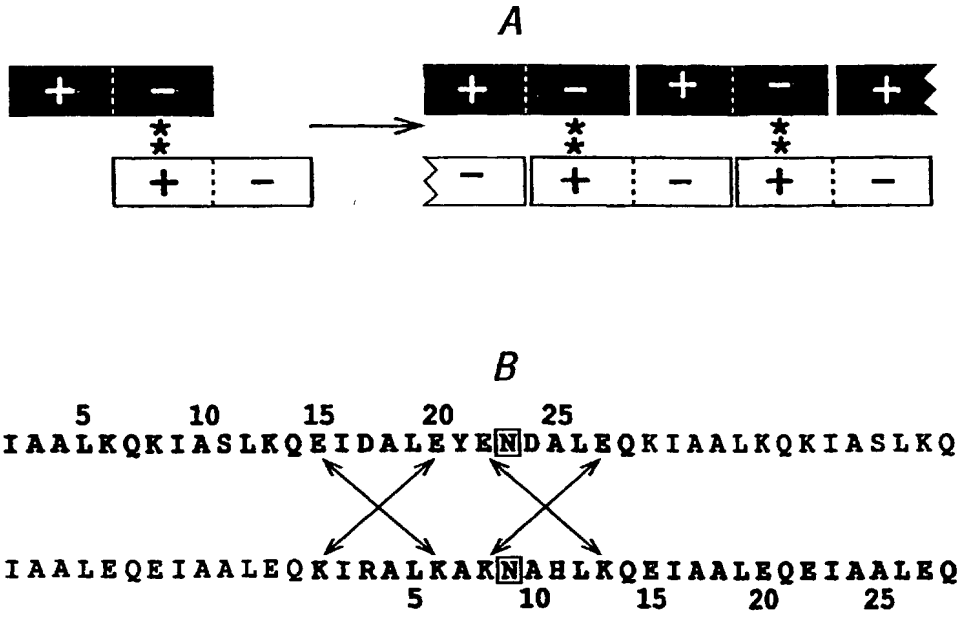
(74) Agents: DEAN, John, Paul et al.; Withers & Rogers, Goldings House, 2 Hays Lane, London SE1 2HW (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: PROTEIN STRUCTURES AND PROTEIN FIBRES



(57) Abstract: This invention relates to protein fibre formation and in particular to methods of producing protein fibres to form a protein structure comprising a plurality of first polypeptide units arranged in a first polypeptide strand and a plurality of second polypeptide units arranged in a second polypeptide strand, the strands preferably forming a coiled coil structure, and in which a first polypeptide unit in the first strand extends beyond a corresponding second polypeptide unit in the second strand in the direction of the strands.

WO 01/21646 A1

Published:

— *With international search report.*

— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

FIG. 1A

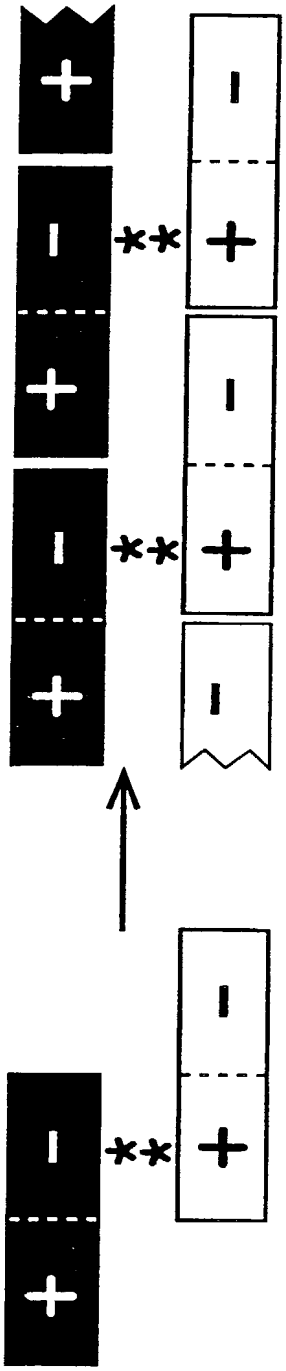


FIG. 1B

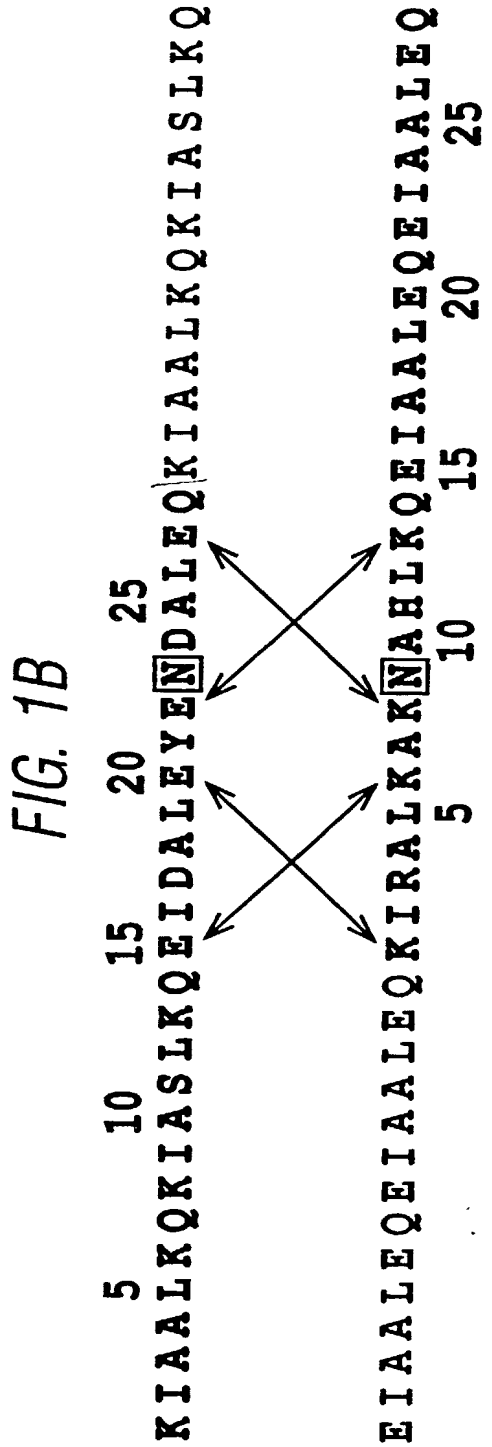

KIAALKQKIASLKQEIDALEYENDALEQKIAALKQKIASLKQ
5 10 15 20 25
EIAALEQEIAALEQKIRALKAKNAHLKQEIAALEQEIAALEQ
5 10 15 20 25

FIG. 1C

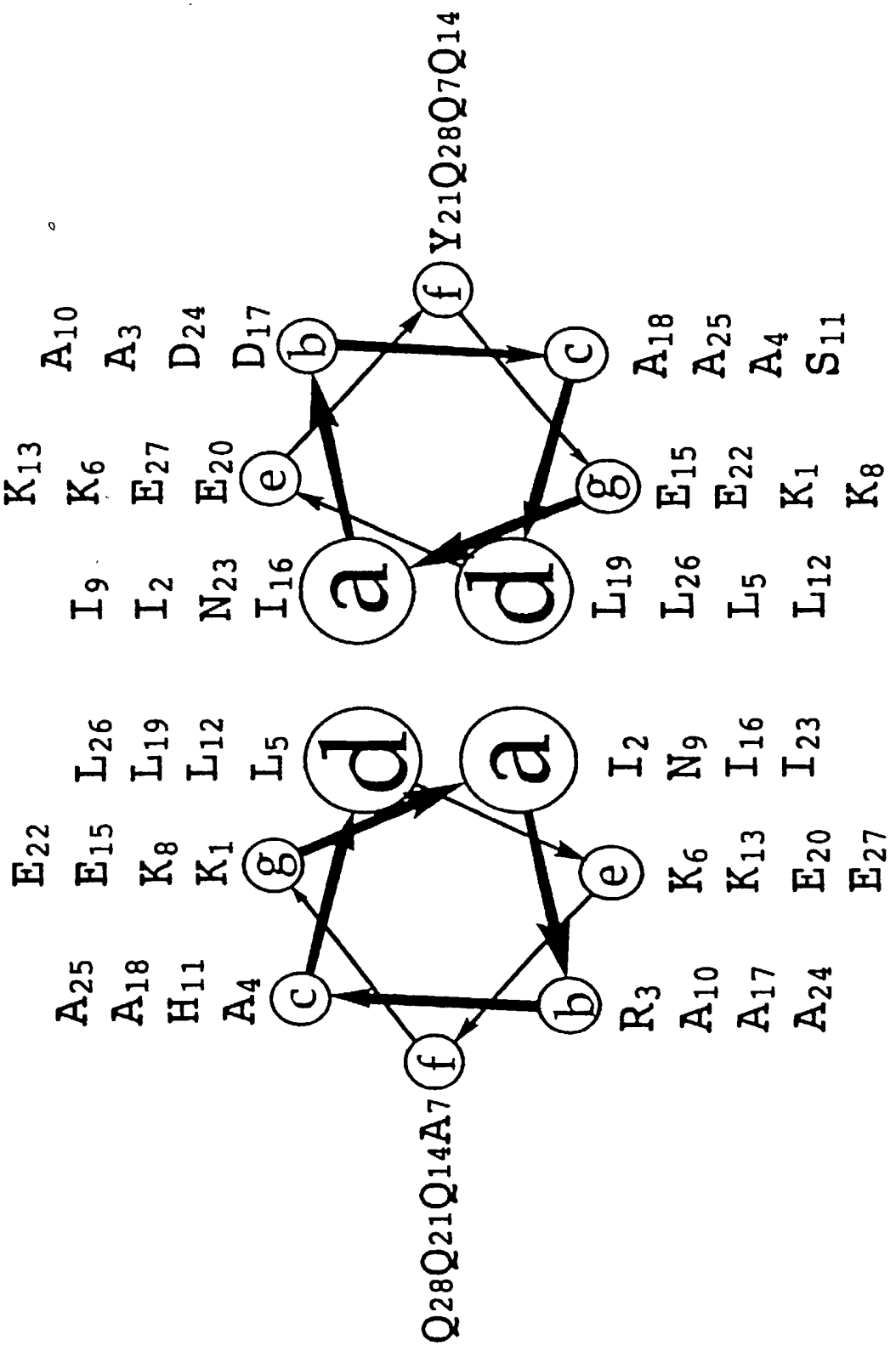

Y21Q28Q7Q14
A10 A3 D24 D17
K13 K6 E27 E20
I9 I2 N23 I16
L19 L26 L5 L12
A18 A25 A4 S11
E22 E15 K8 K1
A25 A18 H11 A4
L26 L19 L12 L5
I2 N9 I16 I23
R3 A10 A17 A24
K6 K13 E20 E27
Q28Q21Q14A7
a
b
c
d
e
f
g

FIG. 2A

FIG. 2B

FIG. 3B

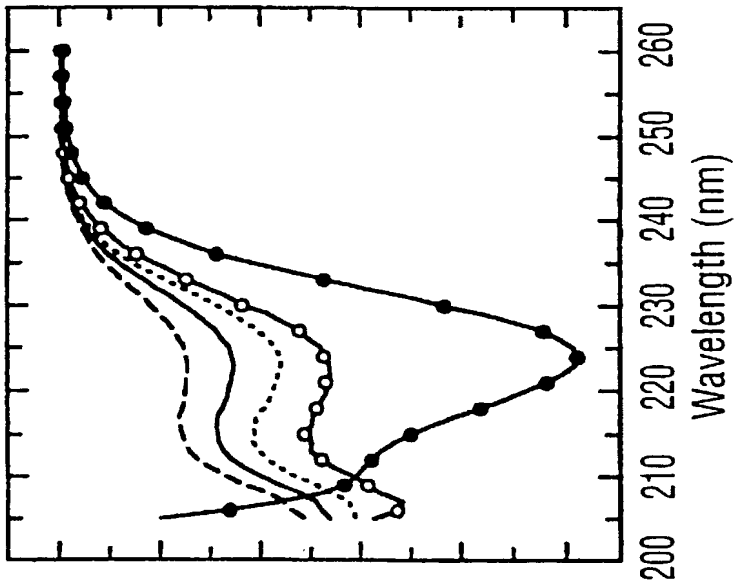

200
210
220
230
240
250
260
Wavelength (nm)

FIG. 3A

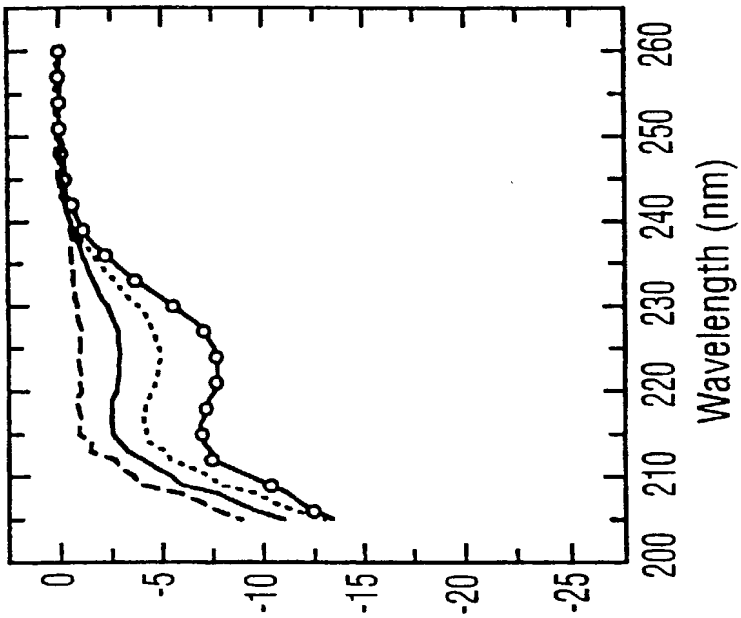

0
-5
-10
-15
-20
-25
Molar Ellipticity
(deg cm2 dmol-res-1 x 10-3)
200
210
220
230
240
250
260
Wavelength (nm)

FIG. 3D

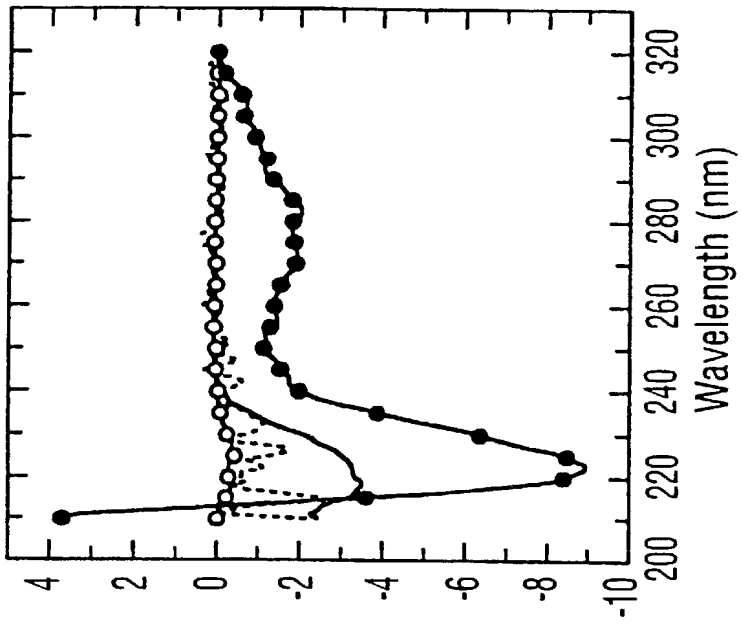

Differential molar extinction coefficient (l mol-res^{-1} cm^{-1})
4
2
0
-2
-4
-6
-8
-10
200
220
240
260
280
300
320
Wavelength (nm)

FIG. 3C

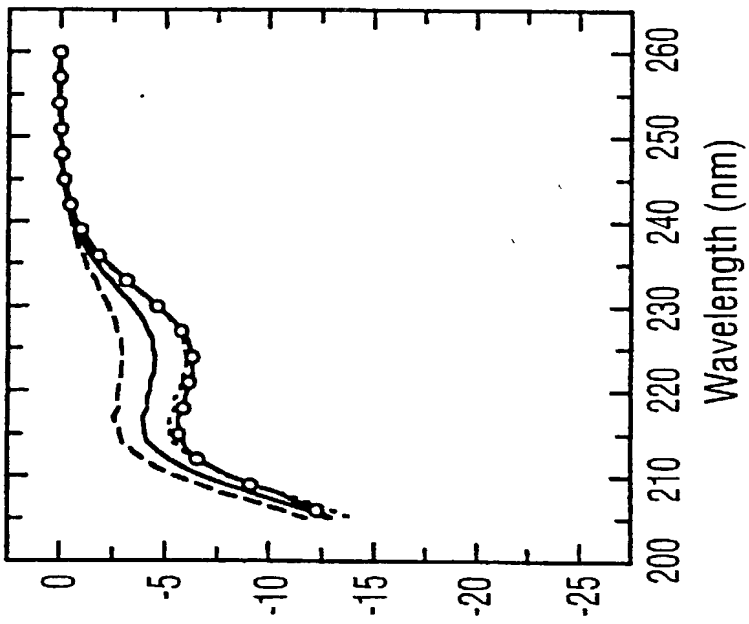

Molar Ellipticity (deg cm^2 dmol-res^{-1} x 10^{-3})
0
-5
-10
-15
-20
-25
200
210
220
230
240
250
260
Wavelength (nm)

FIG. 4B

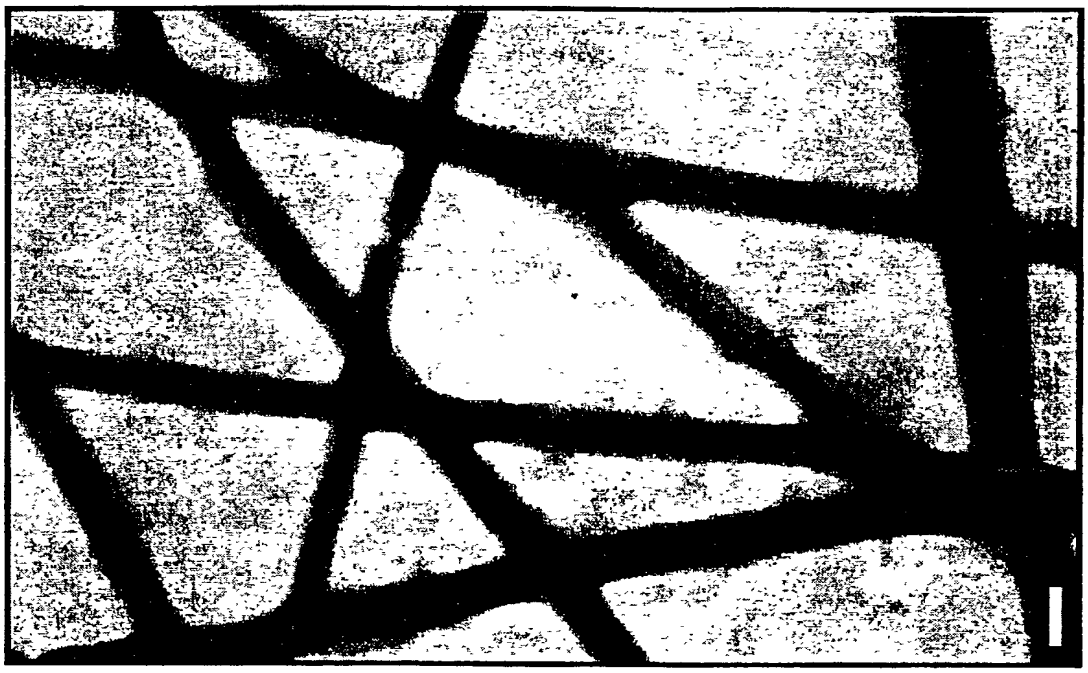


FIG. 4A

FIG. 4E

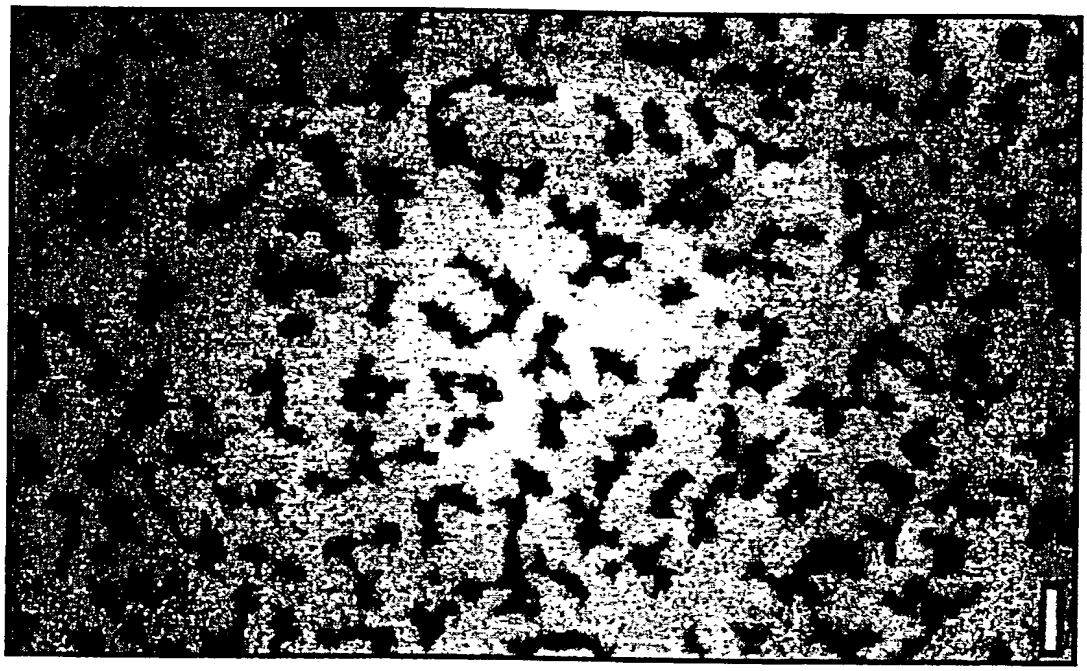
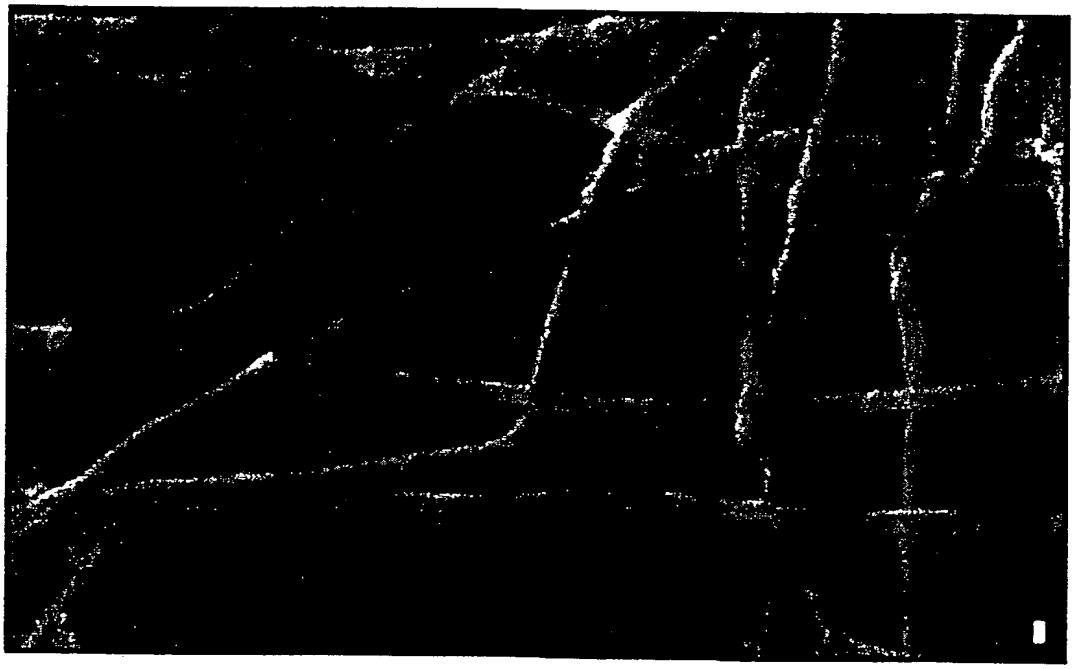


FIG. 4C



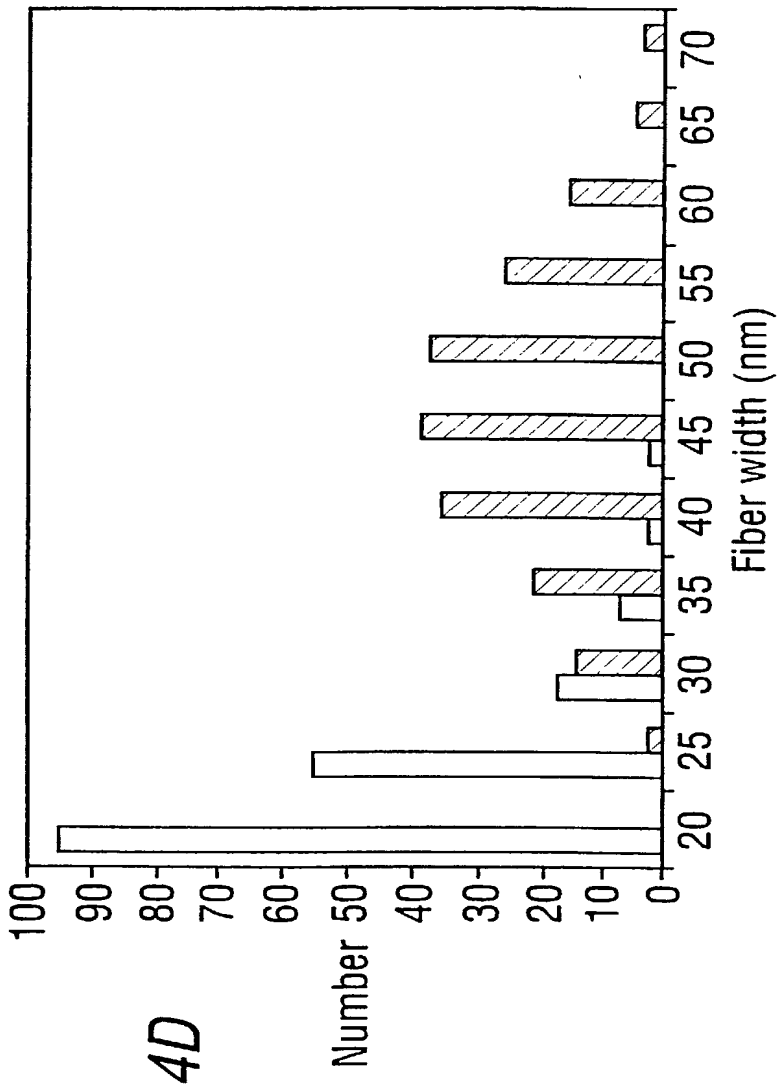
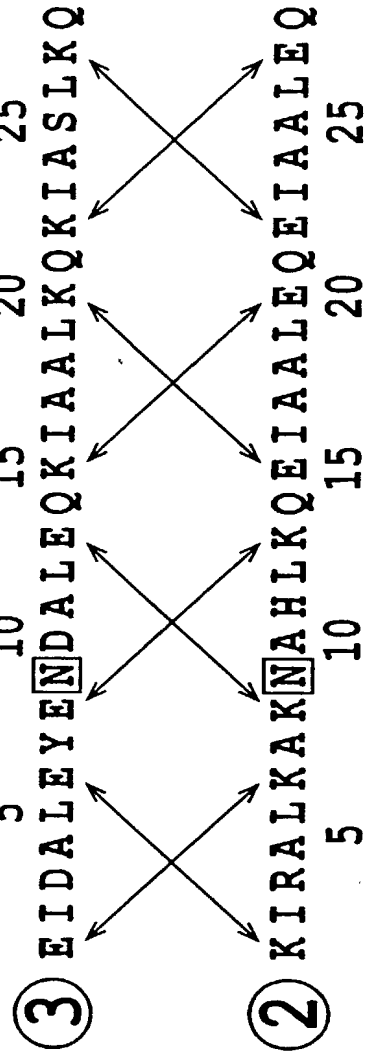

Number
Fiber width (nm)
0
10
20
30
40
50
60
70
80
90
100
20
25
30
35
40
45
50
55
60
65
70

FIG. 4D


③ EIDALEYENDALEQKIAALKQKIASLKQ
② KIRALKAKNAHLKQEIAALEQEIAALEQ
5
10
15
20
25

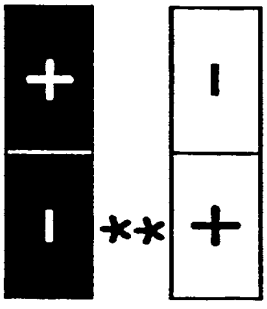

-
+
**
+
-

FIG. 8

FIG. 5

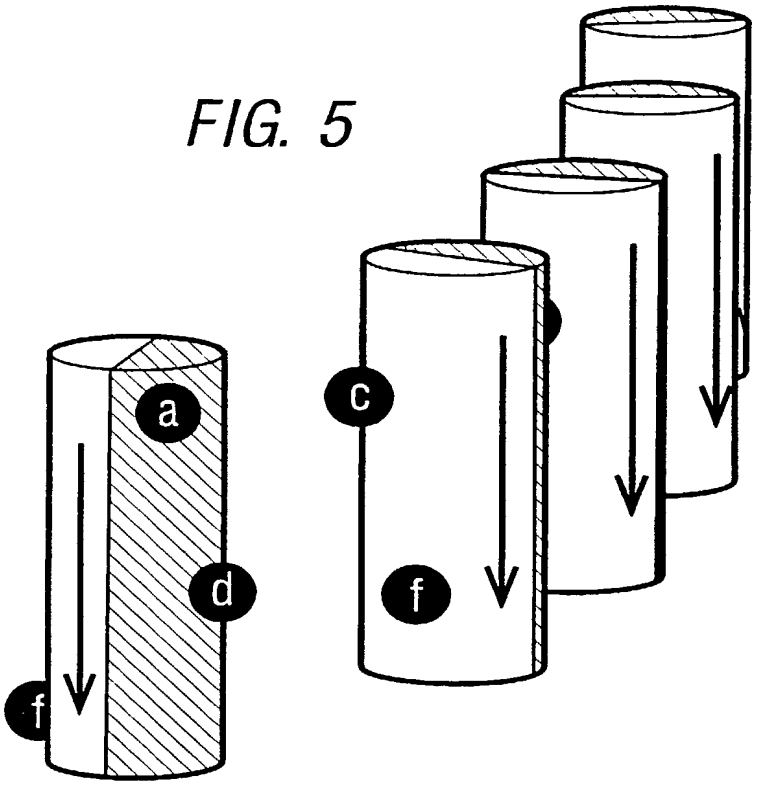
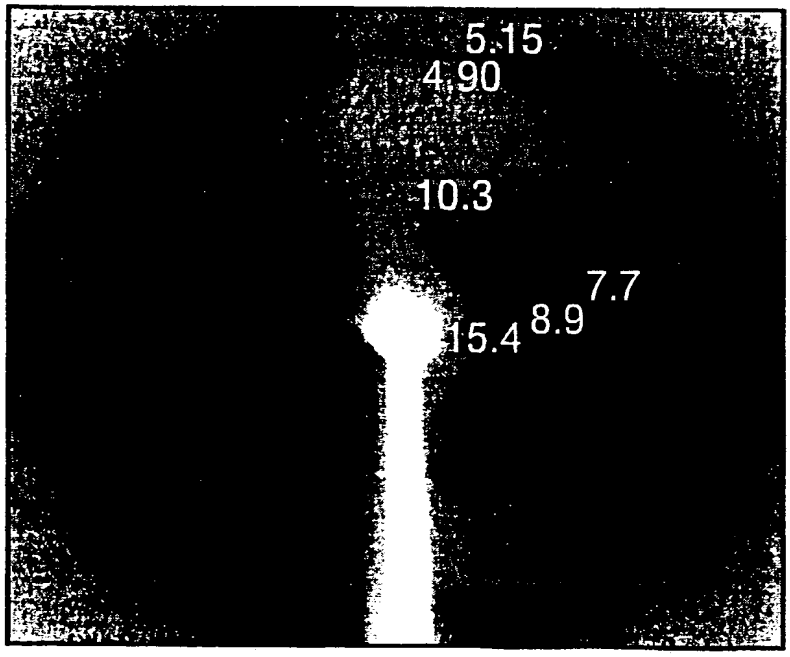


FIG. 6



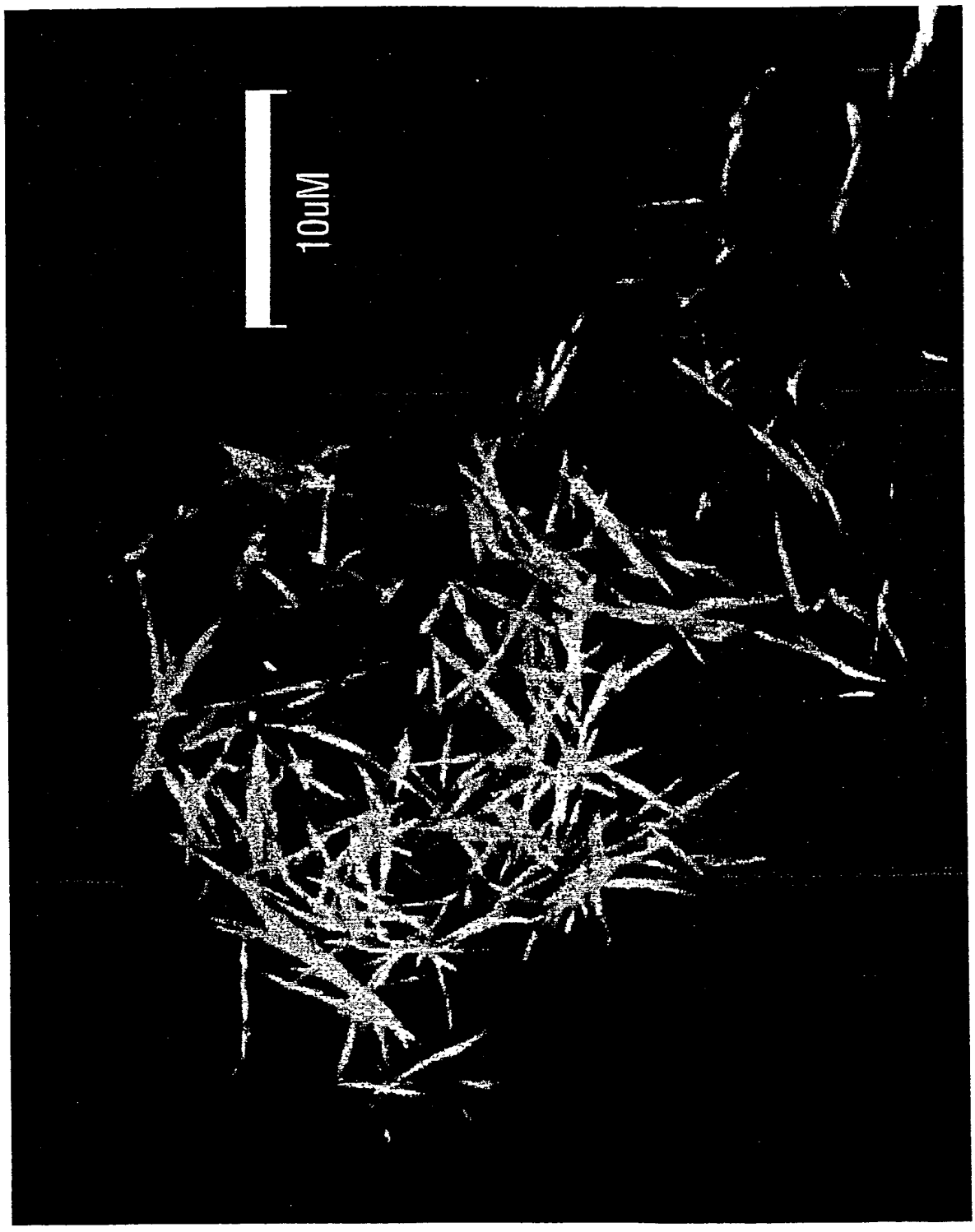


FIG. 7

Banner & Witcoff Ref. No. 00487.00012
Client Ref. No. P100087US-PCT-AGT-jmj

JOINT DECLARATION FOR PATENT APPLICATION

As the below named inventors, we hereby declare that:

Our residence, post office address and citizenship are as stated below next to our names;

We believe we are the original, first and joint inventors of the subject matter which is claimed and for which a patent is sought on the invention entitled PROTEIN STRUCTURES AND PROTEIN FIBRES, the specification of which

- ☐ is attached hereto.
- ☒ was filed on March 18, 2002 as Application Serial Number 10/088,417 and was amended on ___ (if applicable).
- ☒ was filed under the Patent Cooperation Treaty (PCT) and accorded International Application No. PCT/GB00/03576, filed September 18, 2000, and amended on _____ (if any).

We hereby state that we have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

We hereby acknowledge the duty to disclose information which is material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56(a).

Prior Foreign Application(s)

We hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application(s) for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Country	Application No.	Date of Filing (day month year)	Date of Issue (day month year)	Priority Claimed Under 35 U.S.C. §119
Great Britain	9922013.9	17 September 1999		Yes

Prior United States Provisional Application(s)

We hereby claim priority benefits under Title 35, United States Code, §119(e)(1) of any U.S. provisional application listed below:

U.S. Provisional Application No.	Date of Filing (day month year)	Priority Claimed Under 35 U.S.C. §119(e)(1)	

Prior United States Application(s)

We hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, we acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

Application Serial No.	Date of Filing (Day, Month, Year)	Status — Patented, Pending, Abandoned	

10008417 072202

Banner & Witcoff Ref. No. 00487.00012
Client Ref. No. P100087US-PCT-AGT-jmj

Power of Attorney

And we hereby appoint, both jointly and severally, as our attorneys with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected herewith the practitioners at:

Customer Number: 22907 (WDC)

22907

Please address all correspondence and telephone communications to the address and telephone number for this Customer Number.

We hereby declare that all statements made herein of our own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Signature _____ Date 10-06-02
Full Name of First Inventor WOOLFSON Derek N
Family Name First Given Name Second Given Name
Residence Great Britain Citizenship Great Britain
Post Office Address University of Sussex, Falmer, Brighton, Sussex BN1 9QG, Great Britain GBN

Signature _____ Date 21-6-02
Full Name of Second Inventor WALSHAW John
Family Name First Given Name Second Given Name
Residence Great Britain Citizenship Great Britain
Post Office Address University of Sussex, Falmer, Brighton, Sussex BN1 9QG, Great Britain GBN

Signature _____ Date 25th June 2002
Full Name of Third Inventor PANDYA Maya J.
Family Name First Given Name Second Given Name
Residence Great Britain Citizenship Great Britain
Post Office Address University of Sussex, Falmer, Brighton, Sussex BN1 9QG, Great Britain GBN

Signature _____ Date 13-6-02
Full Name of Third Inventor COLYER John
Family Name First Given Name Second Given Name
Residence Great Britain Citizenship Great Britain
Post Office Address Elfordlea, Mill Lane, Bardsey LS17 9AN, Great Britain GBN

#6

SEQUENCE LISTING

<110> Woolfson, Derek
Washaw, John
Pandya, Maya
Colyer, John

<120> PROTEIN STRUCTURES AND PROTEIN FIBRES

<130> 000487.00012

<140> 10/088,417
<141> 2002-03-18

<150> PCT/GB00/03576
<151> 2000-09-18

<150> GB9922013.9
<151> 1999-09-17

<160> 18

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 28
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 1
Lys Ile Ala Ala Leu Lys Gln Lys Ile Ala Ser Leu Lys Gln Glu Ile
1 5 10 15
Asp Ala Leu Glu Tyr Glu Asn Asp Ala Leu Glu Gln
20 25

<210> 2
<211> 28
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 2
Lys Ile Arg Ala Leu Lys Ala Lys Asn Ala His Leu Lys Gln Glu Ile
1 5 10 15
Ala Ala Leu Glu Gln Glu Ile Ala Ala Leu Glu Gln
20 25

<210> 3

<211> 28
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 3
Lys Ile Ala Ala Leu Lys Gln Lys Ile Ala Ala Leu Lys Gln Glu Ile
1 5 10 15
Asp Ala Leu Glu Tyr Glu Asn Asp Ala Leu Glu Gln
20 25

<210> 4
<211> 28
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 4
Lys Ile Arg Ala Leu Lys Trp Lys Asn Ala His Leu Lys Gln Glu Ile
1 5 10 15
Ala Ala Leu Glu Gln Glu Ile Ala Ala Leu Glu Gln
20 25

<210> 5
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 5
Tyr Gly Pro Gly Glu Ile Ala Ala Leu Glu Gln Glu Asn Ala Ala Leu
1 5 10 15
Glu Gln

<210> 6
<211> 28
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 6
Lys Ile Ala Ala Ile Lys Gln Lys Ile Ala Ala Leu Lys Gln Glu Ile
1 5 10 15
Asp Ala Leu Glu Tyr Glu Asn Asp Ala Leu Glu Gln
20 25

<210> 7
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 7
Lys Ile Ala Ala Leu Lys Gln Ile Cys Ile Ala Ala Leu Lys Gln Glu
1 5 10 15
Ile Asp Ala Leu Glu Tyr Glu Asn Asp Ala Leu Glu Gln
20 25

<210> 8
<211> 28
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 8
Lys Ile Ala Ala Leu Lys Gln Lys Ile Ala Ser Leu Lys Gln Glu Ile
1 5 10 15
Asp Ala Leu Glu Tyr Glu Asn Asp Ala Leu Glu Gln
20 25

<210> 9
<211> 28
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 9
Lys Ile Ser Ala Leu Lys Trp Lys Asn Ala Ser Leu Lys Gln Glu Ile
1 5 10 15
Ala Ala Leu Glu Gln Glu Ile Ala Ala Leu Glu Gln
20 25

<210> 10
<211> 28
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 10
Lys Ile Arg Ala Leu Lys Trp Lys Asn Ala His Leu Lys Gln Glu Ile
1 5 10 15
Ala Ala Leu Glu Gln Glu Ile Ala Ala Leu Glu Gln
20 25

<210> 11
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 11
Ile Cys Ile Arg Ala Leu Lys Ala Lys Asn Ala His Leu Lys Gln Glu
1 5 10 15
Ile Ala Ala Leu Glu Gln Glu Ile Ala Ala Leu Glu Gln
20 25

<210> 12
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 12
Ile Thr Ile Arg Ala Leu Lys Cys Lys Asn Ala His Leu Lys Gln Glu
1 5 10 15
Ile Ala Ala Leu Glu Gln Glu Ile Ala Ala Leu Glu Gln
20 25

<210> 13
<211> 28
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 13
Glu Ile Asp Ala Leu Glu Tyr Glu Asn Asp Ala Leu Glu Gln Lys Ile
1 5 10 15
Ala Ala Leu Lys Gln Lys Ile Ala Ser Leu Lys Gln
20 25

<210> 14
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 14
Leu Ala Ala Leu Ala Ala Ala
1 5

<210> 15
<211> 43
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 15
Lys Ile Ala Ala Leu Lys Gln Lys Ile Ala Ser Leu Lys Gln Glu Ile
1 5 10 15
Asp Ala Leu Glu Tyr Glu His His Asp Ala Leu Glu Gln Lys Ile Ala
20 25 30
Ala Leu Lys Gln Lys Ile Ala Ser Leu Lys Gln
35 40

<210> 16
<211> 42
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 16
Glu Ile Ala Ala Leu Glu Gln Glu Ile Ala Ala Leu Glu Gln Lys Ile
1 5 10 15
Arg Ala Leu Lys Ala Lys Gln Ala Lys Leu Lys Gln Glu Ile Ala Ala
20 25 30
Leu Glu Gln Glu Ile Ala Ala Leu Glu Gln
35 40

<210> 17
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 17
Glu Ile Asp Ala Leu Glu Tyr Glu Gln Asn Asp Ala Leu Glu Gln Lys
1 5 10 15
Ile Ala Ala Leu Lys Gln Lys Ile Ala Ser Leu Lys Gln
20 25

<210> 18
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> synthetic peptides

<400> 18

```
Lys Ile Arg Ala Leu Lys Ala Lys Phe Asn Ala His Leu Lys Gln Glu
 1               5                  10                  15
Ile Ala Ala Leu Glu Gln Glu Ile Ala Ala Leu Glu Gln
            20                  25
```

United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division

Application deficiencies found during scanning:

☐ Page(s)_____ of _____ were not present for scanning. (Document title)

☐ Page(s)_____ of _____ were not present for scanning. (Document title)

☐ ***Scanned copy is best available.***

Parts of drawings are very dark.

Lines are crossing the declaration sheets.